PHYLOGENETIC RELATIONSHIPS IN THE GENUS *ALLOGLOSSIDIUM* (DIGENEA: PLAGIORCHIOIDEA): EVOLUTIONARY ORIGINS AND IMPLICATIONS OF CHANGES IN LIFE CYCLE COMPLEXITY

A Dissertation

by

EMILY LOUISE KASL

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Charles D. Criscione
Committee Members,	Mary K. Wicksten
	Adam Jones
	Kirk Winemiller
Head of Department,	Thomas McKnight

August 2016

Major Subject: Biology

ABSTRACT

Parasite life cycle complexity (host species and number needed to complete development) is a key determinant of parasite ecology and evolution because it can influence parasite transmission, gene flow, and mating systems. Many digenean trematodes have a 3-host life cycle, where worm sexual reproduction occurs in a vertebrate definitive host, but some species exhibit a 2-host pattern where sexual maturity occurs in what is typically considered the second intermediate host. In the genus *Alloglossidium*, 2- and 3-host life cycles are present among species and a variety of definitive hosts are used: catfish, crustacean, or leech. Thus, the genus *Alloglossidium* provides a model system to study the evolution of complex life cycles. This dissertation focuses on elucidating the taxonomy and evolutionary relationships within the genus *Alloglossidium* and investigating the consequences of changes in life cycle complexity. I provide the first molecular-based phylogeny, which was subsequently used to estimate the number of origins of changes in life cycle complexity within the genus. The findings of this ancestral reconstruction provide strong evidence for multiple origins of a precocious, 2-host life cycle pattern: an early divergence in those species maturing in leeches and at least 3 separate origins of the precocious life cycles maturing in crustacean hosts. Next, the issues inherent in the recognition and disambiguation of cryptic species are addressed. Based on molecular species delineation, it is recognized that there is a need for taxonomic revision. Thus, an analysis of the subtle morphological variation present among the *Alloglossidium* species from fishes is given.

Additionally, a suite of morphological characters that may be useful in interspecific identification are highlighted (i.e., body size, egg size, extent of the vitellaria, shape and placement of the ovary, and appearance of the cirrus sac). Finally, I discuss how changes in life cycle complexity associated with altered mating systems (i.e., precocious sexual maturation may impact opportunities for outcrossing) influence the evolution of reproductive traits in parasites. I found that sex allocation evolved towards a more female-biased function in populations of the hermaphroditic digenean trematode *Alloglossidium progeneticum* that can precociously reproduce in their second hosts.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Charles Criscione, and my committee members, Dr. Mary Wicksten, Dr. Adam Jones, and Dr. Kirk Winemiller for their guidance and support throughout the course of this research and writing process. I would also like to thank Dr. Bill Font, for teaching me traditional parasitological techniques and taxonomy, and Dr. Norm Dronen, for the use of his equipment in the creation of my many line drawings. Additionally, I would like to acknowledge the collaborative efforts of T. J. Fayton, Dr. Sean Locke, and Dr. David Marcoligese in providing specimens that led to the identification of two new species of *Alloglossidium*.

I would also like to thank those whose help and assistance in conducting the large-scale survey trips (14+ hour driving days, spending nights collecting leeches in the bogs of northern Wisconsin and Minnesota and days trekking through the swamps and streams of the southern U.S., and the many, many hours spent crouched over dissecting scopes in the accommodating atmosphere provided by several motel chains): my fellow graduate students, Mary Gorton-Janicka and Andrew Sakla, and the many undergrads over the past 6 years, including Katie Bass, Katie Peart, Abbie Ince, Nick Stokes, Briana Trejo, and Heather Newkirk.

Thanks also go to my friends and colleagues in the Department of Biology (especially members of the PhD cohort of 2009). My postdoc "advisors," Dr. Jillian Detwiler and Dr. Isabel Caballero, who were invaluable in navigating my research and professional development. And, in particular, Emily Rose and Melissa Giresi, whose

support and encouragement at the most critical moments made it possible to see the light at the end of the tunnel.

Finally, for their unending support and constant encouragement, I would like to thank my parents, Doug and Dana, who in a pinch could also be counted on as copilot and leech collector, and brother Matt, whose inappropriate song choices, amazing woodworking skills, and late night discussions made all the difference.

This research was possible due to the funding of the National Science Foundation (DEB #1145508) and a Willis A. Reid, Jr. Student Research Grant (American Society of Parasitologists).

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Complex life cycles often involve an organism passing through 2 or more distinct ecological and morphological phases during the transition from larvae to adult (Wilbur 1980). In metazoan parasites, life cycle complexity is also associated with 1) the number of hosts needed to complete development, 2) the type of host (e.g., vertebrate or invertebrate) used during a given life history stage, and 3) the site of infection (e.g., encysted in muscle, free in gut tract) within a given host (Roberts and Janovy 2009). Recent theoretical studies focusing on the causes and consequences of life cycle complexity have renewed attention to life cycle evolution among parasites (reviewed in Parker et al. 2015a, 2015b). However, for empirical studies to address questions related to the evolution of life cycle complexity, appropriate study systems are needed. In particular, systems where there is intraspecific variation in life cycle patterns or where life cycle variation is present across closely related species are ideally suited to begin addressing complex life cycle evolution.

Digenean trematodes are notable for having some of the most complex life cycles, typically incorporating both free-living and parasitic developmental stages and always including both asexual and sexual reproduction (Cribb et al. 2003; Olson et al. 2003). Among digeneans life cycles may vary in the number of hosts used to complete development, the type of hosts used as the definitive host (host where the parasite sexually matures), and the number of developmental stages (Cribb et al. 2003). While a

3-host (vertebrate definitive host) life cycle pattern is very common, other variants have precocious development (historically termed progenetic) wherein the parasite attains sexual maturity in what is typically deemed an intermediate host (Cribb et al. 2003; Lefebvre and Poulin 2005). Notable among digenean genera is the genus *Alloglossidium*, which stands out due to a high amount of interspecific life cycle variation. Five life cycle patterns are recognized involving the use of 2- and/or 3-hosts to complete development as well a number of possible definitive host species: catfishes, crustaceans, or leeches. The recognition of this fascinating life cycle variation was due to a succession of species descriptions and life history studies that took place during the 1970s and 80s. The culmination of these studies argues that *Alloglossidium* could serve as a model system with which to study the evolution of parasite development and life cycle complexity (Font 1980; Carney and Brooks 1991; Smythe and Font 2001; Brooks 2003).

However, before the genus *Alloglossidium* can be used as a model, one must first address a number of key issues. First, one must have a strong taxonomic grasp on the genus. Cryptic species (morphologically similar, but genetically distinct) can mislead taxonomy, and coupled with undiscovered species, can create gaps that skew phylogenetic inference. Therefore, wide scale surveys are needed to fully assess the diversity of the genus. Second, a phylogeny is needed to establish the order of transitions involved in changes in life cycle complexity. The success of the genus *Alloglossidium* as a model system is therefore contingent on the ability to overcome these issues. Thus, my dissertation focuses on addressing these 2 problems by

elucidating the taxonomy and evolutionary relationships within the genus *Alloglossidium*. Subsequently, I use this framework to investigate a consequence of a change in life cycle complexity.

My dissertation proceeds as such: in Chapter II, I address the need for an updated phylogeny to assess the evolutionary history of life history traits. I use DNA sequence data to reconstruct the phylogeny of the genus and compare my results to the previous morphological based phylogeny. With the molecular phylogeny, I then assess the number of origins of changes in life cycle complexity. The findings of this ancestral reconstruction provide strong evidence for multiple origins of a precocious, 2-host life cycle pattern. In Chapter III, I address the taxonomic pitfalls and solutions associated with the recent recognition of cryptic diversity (via molecular species delineation) among *Alloglossidium* spp. from catfishes. I provide a discussion of the subtle morphological variation present among these species, with a focus on a subset of 5 species that are in need of taxonomic revision or species description. In Chapter IV, I investigate an evolutionary consequence of a change in life cycle complexity. In particular, I assess the evolution of sex allocation given a change in the mating system that results from a developmental change (precocious development in an intermediate host) in the life cycle. This last study was only made possible via biodiversity surveys that discovered the life history and morphological variation across populations of *Alloglossidium progeneticum*.

CHAPTER II

THE EVOLUTION OF LIFE CYCLE COMPLEXITY IN PARASITES: A MOLECULAR PHYLOGENETIC REVISION OF THE GENUS *ALLOGLOSSIDIUM* REVEALS INDEPENDENT ORIGINS OF PRECOCIOUSNESS

Introduction

Life cycle complexity in parasites involves transmission through one or more intermediate hosts, where only parasite development occurs, prior to infection of a definitive host, where parasite sexual maturation occurs. Interest in this fascinating aspect of parasite biology has received renewed attention due to recent theoretical studies modeling potential selective mechanisms acting on parasite life history, transmission, or mating systems that could lead to the evolution of complex life cycles utilizing trophic transmission (Brown et al. 2001; Choisy et al. 2003; Parker et al. 2003; Parker et al. 2015a, 2015b). These studies reveal that transitions through multiple hosts are not trivial events. For example, the upward incorporation of a new definitive host would require predation on the previous definitive host and higher survivorship and fecundity in the new host. The downward incorporation of an intermediate host requires the potential intermediate host to get infected, be prey of the definitive host, and for transmission to the definitive host to increase (reviewed in Parker et al. 2015a, 2015b). Although there is empirical work on how parasite life history traits have evolved in the context of a given life cycle pattern (Gemmill et al. 1999; Hammerschmidt et al. 2009; Benesh et al. 2012; Benesh et al. 2013), empirical work on the causes and consequences

of changes in complex life cycles themselves would be greatly facilitated by systems where there is either intraspecific or closely related interspecific variation in life cycle patterns is present (Morand et al. 1995). However, to make full use of such life cycle variation, it is paramount to understand the order of events (i.e., the ancestral and derived states), the number of transitions between life cycle patterns, and the traits that may correlate to those transitions. Thus, one will need to first determine the relationships among taxa in order to assess the validity of hypotheses (reconstructions) of the evolutionary pathway of changes in life cycle complexity.

Among the Neodermata (i.e., parasitic flatworms), the Digenea are notable for having the most complex life cycles (Cribb et al. 2003; Olson et al. 2003). A subclass of the Trematoda, digeneans consist of ~18,000 nominal species (Bray et al. 2008), though various estimates suggest that the number of species may in fact be much higher (74,482-126,695 depending on estimation method used; Strona and Fattorini 2014). Across the Digenea, a 3-host life cycle is the most commonly found pattern; however, the most comprehensive trematode phylogeny to date suggests the second host was adopted repeatedly possibly by downward incorporation (Cribb et al. 2003, Olsen et al. 2003). In a 3-host life cycle, miracidia from eggs infect a molluscan first intermediate host. Asexual reproduction within the mollusc leads to the release of larval cercariae that penetrate a second intermediate host (an invertebrate or vertebrate) and therein encyst as metacercariae. Upon ingestion of the second intermediate host by a definitive host, the parasite undergoes maturation and obligate sexual reproduction. The life cycle

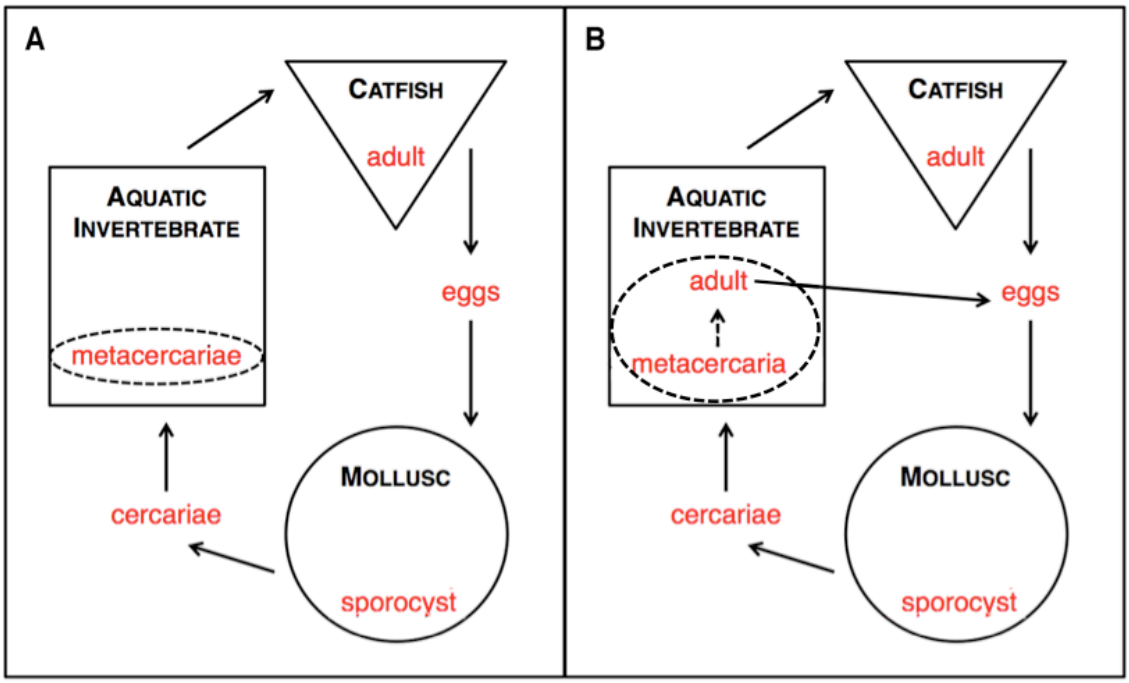

Figure 1. Diagram of two *Alloglossidium* life cycles: (**A**) obligate 3-host and (**B**) facultative 2- or 3-host. In the obligate 3-host life cycle (A) an egg that contains a larval miracidium is ingested by a molluscan first host, wherein the parasite undergoes asexual reproduction. Larval cercariae leave the first host and penetrate an invertebrate second intermediate host. The worms remain encysted as metacercariae in the second host until ingestion by a definitive (ictalurid) host, where the parasite becomes sexually mature. In contrast, in a facultative precocious life cycle (**B**), the parasite becomes sexually mature while still encysted in a crayfish second host (i.e., a fish is not needed). However, as encysted worms can still be trophically transmitted to an ictalurid host, this life history is considered a facultative 2- or 3-host life cycle.

continues with the release of eggs, which contain the miracidia, into the environment (Fig. 1A).

Groups of closely related (intra or interspecific) digeneans that show variation in their life cycle patterns can be useful for studying the evolution of complex life cycles. One interesting digenean life cycle variant involves precocious development (historically termed progenetic) such that there is early onset of sexual maturation in what is typically deemed an intermediate host (Lefebvre and Poulin 2005). Several hypotheses have been put forth to explain precocious development: the need to mitigate disruptions to transmission caused by environmental instability, lack of hosts, or developmental time (due to long term encystment in an intermediate host) (Poulin and Cribb, 2002; Lefebrve and Poulin 2005a). The most recent wide-scale review of precociousness in digeneans identified 79 digenean species (spanning 50 genera and 24 families) as exhibiting precocious development (Lefebvre and Poulin 2005). A disproportionately high amount of these 79 species (roughly 15%) were associated with a single digenean group: the genus *Alloglossidium* Simer, 1929.

At the time of this writing, there were 18 nominal species within the genus *Alloglossidium* and among these species, there are 5 recognized life cycle variations. The obligate 3-host life cycle reflects the common digenean pattern wherein the definitive host is a vertebrate, an ictalurid catfish (Fig. 1A). Although this pattern has only been experimentally determined in *Alloglossidium corti* (Crawford 1937), it is assumed to be the life cycle of other *Alloglossidium* species that have only been reported as adults from catfishes: *Alloglossidium kenti*, *Alloglossidium geminum*, *Alloglossidium*

fonti, and *Alloglossidium floridense* (Table 1; Simer 1929; Mueller 1930; Tkach and Mills 2011; Kasl et. al. 2014). Furthermore, sequence data have been used to delineate intraspecific life cycle variation in *Alloglossidium progeneticum*. Some populations obligately use 3 hosts to complete development (Table 1; Kasl et al. 2015) whereas a few populations have facultative precocious life cycles involving the sexual maturation of the parasite while still encysted in a crayfish second host (Table 1; Fig. 1B; Sullivan and Heard 1969; Font and Corkum 1975). These encysted worms can still be trophically transmitted to an ictalurid host, thereby facultatively using 2- or 3-hosts. The facultative life cycle pattern has only been reported from a few populations of *A. progeneticum* in Georgia (see Chapter IV).

The majority of *Alloglossidium* species are reported as utilizing an obligate 2-host life cycle to complete development. I further classify these precocious life cycles by the definitive host (i.e., leech or freshwater crustacean) and the presence or absence of an encysted metacercarial developmental stage. Species using haemopid or macrobdellid leeches as definitive hosts transition through all developmental stages: cercariae released from a molluscan first host penetrate the leech, encysting as metacercariae in the crop. Upon excysting, worms migrate within the same host individual to the gut lumen, where they sexually mature (Fig. 2A; Corkum and Beckerdite 1975). Seven species are known to exhibit this life history: *Alloglossidium hirudicola*, *Alloglossidium macrobdellensis*, *Alloglossidium richardsoni*, *Alloglossidium hamrumi*, *Alloglossidium turnbulli*, *Alloglossidium schmidti*, and *Alloglossidium demshini* (Schmidt and Chaloupka 1969; Beckerdite and Corkum 1974; Fish and Vande

Table 1. Sample ID, species name (based on morphological identification), definitive host name, definitive host type (C = crustacean; F = ictalurid catfish; L = leech), number of hosts used to complete life cycle, presence of an encysted metacercarial stage (0 = absent; 1 = present), collection locality, latitude (N), longitude (W).

ID	Species	Definitive Host	Host Type	# of Hosts	Encyst. Meta.	Collection locality	Latitude (N)	Longitude (W)
AN01	*A. anomophagis*	*Daphnia obtusa* *	C	2	1	Houston Coastal Center, TX ^	29° 43.34394'	95° 31.1331'
AN02	*A. anomophagis*	*Daphnia obtusa* *	C	2	1	Houston Coastal Center, TX ^	29° 43.34394'	95° 31.1331'
CA01	*A. cardicolum*	*Procambarus acutus* *	C	2	0	Walnut Rd, MS	32° 40.354'	89° 43.907'
CA02	*A. cardicolum*	*Procambarus acutus* *	C	2	0	Rosedale, LA ^	30° 27.727'	91° 25.023'
DO01	*A. dolandi*	*Procambarus acutissimus*	C	2	0	Jackson Br. @ Williams Store Rd, GA ^	32° 43.095'	81° 30.133'
DO02	*A. dolandi*	*Procambarus epicyrtus* *	C	2	0	Jackson Br. @ Newington Hwy, GA ^	32° 42.930'	81° 29.362'
GR01	*A. greeri*	*Cambarellus schudfeldti* *	C	2	0	Montrose, LA ^	31° 34.55'	92° 59.85'
GR02	*A. greeri*	*Cambarellus schudfeldti* *	C	2	0	Vatcherie, LA	29° 54.729'	90° 43.713'
GR03	*A. greeri*	*Cambarellus schudfeldti* *	C	2	0	Rockwood, IL	37° 43.546692'	89° 27.809653'
RE01	*A. renale*	*Palaemonetes kadiakensis* *	C	2	0	Olustee Creek, AL	31° 56.65'	86° 7.133333'
RE02	*A. renale*	*Palaemonetes kadiakensis* *	C	2	0	Choctaw Rd, LA	29° 51.478'	90° 45.281'
PR01	*A. progeneticum*	*Ameiuras natalis*	F	3	1	Crooked Creek, AR	36° 14.116'	92° 42.763'
PR02	*A. progeneticum*	*Procambarus spiculifer* *	C/F	2/3	1	Calls Creek, GA ^	33° 53.310'	83° 22.918'
PR03	*A. progeneticum*	*Ameiurus natalis*	F	3	1	Gus Engeling WMA, TX	31° 55.687'	95° 53.279'
CO01	*A. corti*	*Ameiurus melas*	F	3	1	Whisky Bay, LA	30° 23.479'	91° 20.826'
CO02	*A. corti*	*Ameiuras natalis*	F	3	1	Greenwood, LA	32° 28.025'	93° 58.905'
CO03	*A. corti*	*Ictalurus punctatus*	F	3	1	Brazos River, Bryan, Texas	30° 37.705167'	96° 32.659333'
FO01	*A. fonti*	*Ameiurus melas* *	F	3	1	Trout Brook, VT	43° 53.36'	72° 41.426667'
FO02	*A. fonti*	*Ameiurus melas* *	F	3	1	Sixmile Creek, WI	43° 8.430613'	89° 25.553496'
FO03	*A. fonti*	*Ameiurus melas*	F	3	1	Greenwood, LA	32° 28.025'	93° 58.905'
FO04	*A. fonti*	*Ameiurus melas*	F	3	1	Trinity River, TX	30° 16.578'	94° 47.763'
MS01	*A. kenti* MS	*Ictalurus punctatus* *	F	3	1	Money Bayou, Money, MS ^	33° 39.493'	90° 12.528'
MS02	*A. kenti* MS	*Ictalurus furcatus*	F	3	1	Tallahatchie River @ Money, MS (N)	33° 43.120'	90° 13.174'
MS03	*A. kenti* MS	*Ictalurus furcatus*	F	3	1	Tallahatchie River @ Money, MS (S)	33° 35.6874'	90° 11.6990'
TK01	*A. kenti* TK	*Ictalurus punctatus* *	F	3	1	Little Brazos, TX	30° 38.485'	96° 31.222333'
TK02	*A. kenti* TK	*Ictalurus punctatus* *	F	3	1	Little Brazos, TX	30° 38.485'	96° 31.222333'
TK03	*A. kenti* TK	*Ictalurus punctatus* *	F	3	1	Crooked Creek, AR	36° 14.116'	92° 42.763'
TK04	*A. kenti* TK	*Ictalurus punctatus* *	F	3	1	Oneida Lake, NY	43° 11.005'	75° 59.636'
TK05	*A. kenti* TK	*Ictalurus punctatus* *	F	3	1	Oneida Lake, NY	43° 11.005'	75° 59.636'
FL01	*A. floridense*	*Noturus gyrinus*	F	3	1	Spring Run, FL ^	29° 51.533833'	82° 44.094'
FL02	*A. floridense*	*Noturus leptacanthus* *	F	3	1	Spring Run, FL ^	29° 51.533833'	82° 44.094'
GE01	*A. geminum*	*Ameiurus nebulosus* *	F	3	1	Oneida Lake, NY ^	43° 11.005'	75° 59.636'
GE02	*A. geminum*	*Ameiurus nebulosus* *	F	3	1	Oneida Lake, NY ^	43° 11.005'	75° 59.636'
CD01	*Alloglossidium n. sp. 1*	*Ictalurus punctatus* *	F	3	1	St. Lawrence River, Quebec, Canada ^	45° 18.96'	73° 52.74'
CD02	*Alloglossidium n. sp. 1*	*Ictalurus punctatus* *	F	3	1	St. Lawrence River, Quebec, Canada ^	45° 18.96'	73° 52.74'

* = indicates type host ^ indicates type locality

Table 1. Continued.

ID	Species	Definitive Host	Host Type	# of Hosts	Encyst. Meta.	Collection locality	Latitude (N)	Longitude (W)
RI01	*A. richardsoni*	*Haemopis plumbea* *	L	2	1	Zippel Bay, Lake of the Woods, MN	48° 52.369'	94° 51.549'
RI02	*A. richardsoni*	*Haemopis marmorata*	L	2	1	Falcon Lake, Canada	49° 40.701'	95° 18.867'
MA01	*A. macrobdellensis*	*Macrobdella decora*	L	2	1	Boot Bog, WI	45° 5.582'	91° 20.178'
MA02	*A. macrobdellensis*	*Macrobdella ditetra* *	L	2	1	Whisky Bay, LA ^	30° 23.479'	91° 20.826'
HA01	*A. hamrumi*	*Macrobdella decora* *	L	2	1	Storrs, U. Connecticut, CT	41° 49.0495'	72° 15.542833'
HA02	*A. hamrumi*	*Macrobdella decora* *	L	2	1	Storrs, U. Connecticut, CT	41° 49.0495'	72° 15.542833'
HI01	*A. hirudicola*	*Haemopis plumbea*	L	2	1	Zippel Bay, Lake of the Woods, MN	48° 52.369'	94° 51.549'
HI02	*A. hirudicola*	*Macrobdella decora*	L	2	1	Zippel Bay, Lake of the Woods, MN	48° 52.369'	94° 51.549'
TU01	*A. turnbulli*	*Haemopis grandis* *	L	2	1	Zippel Bay, Lake of the Woods, MN ^	48° 52.369'	94° 51.549'
TU02	*A. turnbulli*	*Haemopis grandis* *	L	2	1	Falcon Lake, Canada	49° 40.701'	95° 18.867'
SC01	*A. schmidti*	*Haemopis grandis* *	L	2	1	Falcon Lake, Canada ^	49° 40.701'	95° 18.867'
SC02	*A. schmidti*	*Haemopis grandis* *	L	2	1	Lake Kabetogama, MN	48° 26.839'	93° 2.951'
GU01	*Alloglossidium n. sp. 2*	*Macrobdella ditetra* *	L	2	1	Gus Engeling WMA, TX ^	31° 55.687'	95° 53.279'
GU02	*Alloglossidium n. sp. 2*	*Macrobdella ditetra* *	L	2	1	Gus Engeling WMA, TX ^	31° 55.687'	95° 53.279'

* = indicates type host ^ indicates type locality

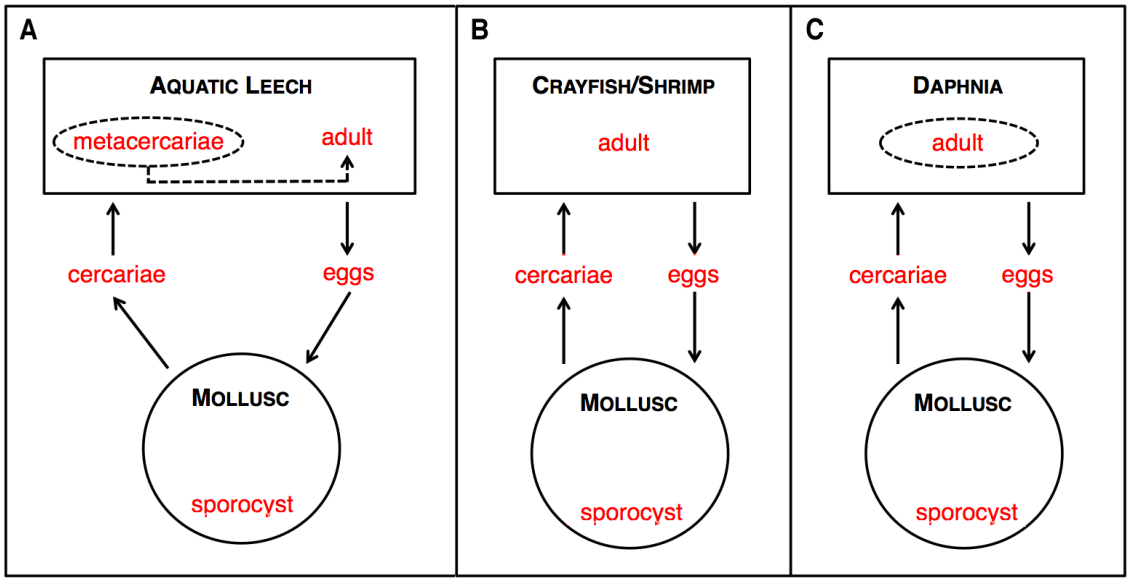

Figure 2. Diagram of *Alloglossidium* life cycles of species with obligate precocious development. Species utilizing leech final hosts (**A**) maintain all life history stages (i.e., the encysted metacercarial stage) with the worm migrating within the host to complete development. For those species that use a crustacean final host, the life cycles can be further distinguished by the presence or absence of an encysted metacercariae. Species reported from the antennal glands of crayfish and shrimp (**B**) have never been reported to have a cyst stage. Instead immature worms and adults have been reported freely in the antennal glands of their respective hosts. In contrast, the final life cycle pattern (**C**) involves sexual maturation of the worm while encysted in the body cavity of *Daphnia*.

Vusse 1976; Neumann and Vande Vusse 1976; Timmers 1979; Tkach et al. 2013). The remaining precocious life cycles involve the penetration of cercariae into a freshwater crustacean serving as the definitive host. Species of *Alloglossidium* reported from grass shrimp (*Alloglossidium renale*) and those from crayfishes belonging to the genera *Procambarus* (*Alloglossidium cardicolum*, *Alloglossidium dolandi*) or *Cambarellus* (*Alloglossidum greeri*) are found freely within the antennal glands of their respective hosts and have lost the encysted metacercarial stage (Fig. 2B; Font and Corkum 1975; Corkum and Turner 1977; Turner and McKeever 1993; Font 1994). Finally, *Alloglossidium anomophagis* reaches sexual maturity within a cyst in the body cavity of *Daphnia* spp. (Fig. 2C; Poinar et al. 1995).

The recognition of these fascinating life cycle variants was due to a quick succession of species descriptions and life history studies conducted on *Alloglossidium* during the 1970s and 80s. Taken together, these papers highlight the genus as a potential model system with which to study questions related to the "evo-devo" (i.e., developmental constraints) and integrative biological (i.e., the origin, stasis, and diversification) components of complex life cycles (Brooks 2003). Indeed, papers began to address the evolutionary history of the group relatively quickly after the initial species discoveries in the 1970s. The first two of these studies used logical deduction to hypothesize the evolutionary history of life cycle changes in the genus. Font (1980) suggested that the ancestral life cycle was a 3-host pattern (fish definitive host) because this was the most common pattern observed among digeneans within the order Plagiorchiida (Cribb et al. 2003). He deduced there was then a split leading to the

speciation of *Alloglossidium* maturing in crustacean hosts and those maturing in leech hosts. Within the crustacean lineage, Font suggested that the obligate 2-host pattern came about through a progressive transition via the facultative precocious pattern. Riggs and Ulmer (1980) postulated that the obligate 2-host pattern in leeches was the ancestral condition due to the fact that leeches predated catfishes in the fossil record. They argued that the similarity between the gut lumen of leeches and the intestines of catfishes were similar enough to enable a host-switching event from leeches to catfishes. However, no species of *Alloglossidium* associated with crustaceans were considered in the Riggs and Ulmer hypothesis.

The first phylogenetic based hypothesis for life cycle change in the genus *Alloglossidium* came from a phylogeny based on 14 morphological traits across 9 species (Carney and Brooks 1991). Smythe and Font (2001) extended this work by including 4 additional morphological characters and 3 life history traits (i.e., 21 total traits) and placing *Alloglosoides dolandi*, *Alloglossoides cardicola*, and *Hirudicolotrema richardsoni* into the genus *Alloglossidium,* bringing the total nominal species up to 14 (Fig. 3; but see Brooks 2003 for methodological corrections). However, a caveat of Smythe and Font's (2001) approach was the use of life history traits to reconstruct the phylogeny that was ultimately used to assess the ancestral states of those traits themselves. Ideally, to test hypotheses related to trait evolution, the phylogeny would be constructed from data independent of the traits themselves.

Here I provide the first molecular-based phylogeny of species in the genus *Alloglossidium* with an emphasis on elucidating the patterns essential to study the

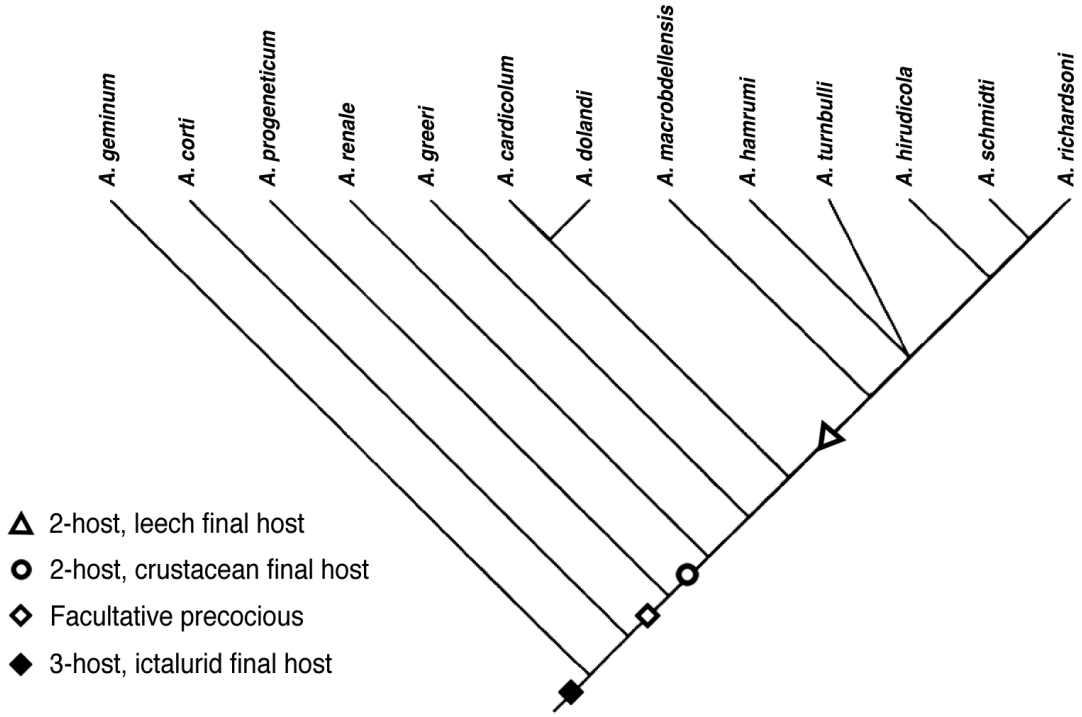

Figure 3. Morphological and life history-based phylogenetic tree depicting the relationships within the genus *Alloglossidium* (adapted from Smythe and Font 2001; see also Brooks 2003). Assumptions and caveats associated with this phylogeny are discussed in the text.

evolution of life cycle complexity, i.e., the ancestral and derived life cycle patterns and the number of independent transitions of life cycle changes. Comparison of the molecular phylogeny to previously proposed hypotheses for the origin of life cycle complexity in the genus *Alloglossidium* indicate the number of independent transitions in life cycle patterns was previously underestimated using the morphological phylogeny. Importantly, this phylogeny establishes a necessary framework with which to assess future evolutionary hypotheses concerning changes in life cycle complexity.

Methods

Taxon Sampling and Outgroup Selection

Sampling of *Alloglossidium* species was conducted as part of a large-scale biodiversity survey of the eastern two-thirds of the United States and southern Canada. All nominal species, with the exception of *Alloglossidium demshini* (described post-survey; Tkach et al. 2013) were collected (Table 1), killed with hot water or heat-fixed underneath a coverslip, and stored in 70% or 95% ethanol for further molecular work. All specimens were identified using the most complete morphological descriptions available in the literature and, when possible, compared to molecular sequences deposited on GenBank (Tkach and Mills 2011: JF440783.1, *A. corti*; JF440765.1and JF44767.1, *A. fonti*; JF440771.1, *A. geminum*; JF440808.1, *A. kenti*; Kasl et al. 2014: KC812276.1, *A. floridense*) for further delimitation. In order to definitively connect molecular sequence data to morphological descriptions, efforts were made to obtain specimens from the type hosts and type localities. Though, it should be noted that some of these locations have been lost due to extirpation of hosts at the type locality (e.g., the

loss of the madtoms hosts used by *A. corti* in Lake Mendota, WI; Lyons 1989), habitat destruction (e.g., *A. renale*; personal observation), or even a lack of type locality (e.g., *A. hirudicola* was originally described from leeches purchased from a biological supply company). All individual specimens used in the phylogenetic analyses were collected from different host individuals. When possible, specimens were also chosen to represent the known geographical ranges (Table 1). In addition, new lineages, some of which likely represent new species, found during the surveys were also included. Two outgroup taxa were used. *Brachycladium goliath* (GenBank # KR703279; Briscoe et al. 2016) was chosen because both the mitochondrial and ribosomal regions of used in our analyses were available on Genbank. Also, this species belongs to the superfamily Allocreadioidea, which is sister to Plagiorchioidea, the superfamily that *Alloglossidium* is supposed to belong. *Paramacroderoides echinus*, from *Lepisosteus osseus* (longnose gar) in the Little Brazos River, TX (TX; 30° 38.485 N 96° 31.222333' W), was chosen due to the accessibility of obtaining specimens for molecular use and the current placements (based on morphology) of both *Alloglossidium* and *Paramacroderoides* within family Macroderoididae McMullen, 1937 (Bray et al. 2008). I note, however, there has been debate over the validity of this family, with a number of attempts made to break up the group long considered to be a taxonomic "catch-all" (Bray et. al 2008). Furthermore, molecular phylogenetic analyses of the 28s ribosomal DNA region in digeneans, using sequences published on GenBank, suggests that the family Macroderoididae may not by monophyletic (E. Kasl et al., unpubl. data). Though I recognize the state of these relationships are in dispute, I defer to the Macroderoididae

classification in choosing the best candidate (i.e., closely related) species for use as the outgroup in these analyses.

Extraction, DNA Amplification, and Sequencing

DNA was extracted by placing individual worms in 200 µl of 5% chelex containing 0.2 mg/ml of proteinase K, incubated for 2 hr at 56°C and boiled at 100°C for 8 min. Two unlinked markers were used for identification and analysis of phylogenetic relationships: 1) a 2591 base pair (2555 bp without gaps) ribosomal DNA (rDNA) fragment spanning from the 3' end of the 18s (309 bp), through the first internal transcribed spacer region (ITS1; ranging from 508-534 bp), the 5.8s (156 bp), the second internal transcribed spacer region (ITS2; ranging from 239-249 bp), to a 1351 bp fragment of the 5' end of the 28s gene (domains D1-D3), and 2) a 676 base pair region of the mitochondrial NADH-dehydrogenase subunit 1 gene (ND1). Polymerase chain reaction (PCR) amplifications were performed using 25 µl reactions. The 18s-ITS1 portion of the nuclear ribosomal sequence (s18 forward primer, 5.8s1 reverse primer; Table 2) was obtained using a reaction consisting of 3 µl of template extract, 16.25 µl water, 2.5 µl 10x buffer, 1.5 µl $MgCl_2$ [25 mM], 0.5 µl dNTP [10 mM/each], 0.5 µl of each primer [20 µM], and 0.25 µl of Taq polymerase (Omega Bio-Tek Inc., Norcross, GA) [5 units/µl], and a thermocycling profile of 95°C for 3 min, once; 94°C for 45 sec, 60°C for 30 sec, 72°C for 60 sec, 35 times; 72°C for 7 min, once. The ITS1-5.8s-ITS2-28s portion of the nuclear ribosomal sequence (CC41 forward primer, 1500R reverse primer; Table 2; thermocycler profile described in Olson et al. 2003) and the ND1 region

Table 2. Primers used for PCRs and sequencing of *Alloglossidium* species. Gene location, primer ID, sequence, reference, orientation and primer function are included. For thermocycler conditions see text.

Location	Primer ID	Primer Sequence 5'-3'	Reference	Orientation: F	PCR (P) or Sequencing
Ribosomal					
18s	s18	TAACAGGTCTGTGATGCC	Jousson et al. 2000	F	P/S
5.8s	5.8s1	GCTGCGCTCTTCATCGACA	Jousson et al. 2000	R	P
28s	1500R	GCTATCCTGAGGGAAACTTCG	Olson 2003	R	P/S
18s	CC41	GATTGAATGGTTTAGCAAGG	Kasl et al. 2014	F	S
5.8s	CC48	TGTCGATGAAGAGTGCAGC	Kasl et al. 2014	F	S
28s	400R	GCAGCTTGACTACACCCG	Olson et al. 2003	R	S
28s	900F	CCGTCTTGAAACACGGACCAAG	Olson et al. 2003	F	S
Mitochondrial					
ND1	MB352	CGTAAGGGKCCTAAYAAG	Criscione and Blouin 2004	F	P/S
ND2	MB411	CATATGATGTTRTCTTCTAG	Criscione and Blouin 2004	F	P
ASN-tRNA	CC28	CWTCTCAARGTTAACAGCCT	this study	R	P
ND3	CC29	GACANAAACCCACACTCAAA	this study	R	P
ND1	CC57	CCCATAATCTATGTGTGCTAAC	Kasl et al. 2015	R	S
ND1	CC79	ATCTTGCTCTTATTGTGTG	this study	F	S

(MB352 forward primer, CC28 reverse primer; MB411 forward primer, CC29 reverse primer; Table 2; thermocycler profile described in Criscione and Blouin 2004) were the obtained using the same reaction recipe with the exception that 15.25 µl water and 2.5 µl $MgCl_2$ [25 mM] were used. If the above protocol was not successful in obtaining ND1 sequences (particularly in the case of the leech-associated species and *A. floridense* and *A. dolandi*) a combination of gel extractions, using the Ultra Clean Gel Spin DNA extraction kit (MO BIO Laboratories, Inc., Solana Beach, California), and alternate sequencing primers were used (Table 2). PCR products were purified using the Ultra Clean PCR clean-up Kit (MO BIO Laboratories, Inc., Solana Beach, California) and sent to the DNA Analysis Facility on Science Hill at Yale University (New Haven, Connecticut) for sequencing.

Phylogenetic Inference Methods

At least 2 individuals per species (as defined by morphological identifications) were used to assess the phylogenetic relationships within the genus *Alloglossidium*. These individual specimens were obtained from unique hosts and, when possible, were chosen to reflect some of the known host diversity (different host species) and/or geographic variation (different locations) associated with each species. Contiguous sequences from individuals were assembled and manually edited using Sequencher™ (GeneCodes Corp., ver. 4.1.4). Both nuclear and mitochondrial sequences were aligned using Clustal W with subsequent adjustment by eye within the BioEdit program, version 7.1.8 (Hall, 1999). To infer phylogenetic relationships, identical sequences were condensed into a single taxonomic unit and used in the analysis of 3 data sets: rDNA-

only (18s-ITS1-5.8s-ITS2-28s), ND1-only, and a concatenated rDNA-ND1 set. The ITS1 and ITS2 regions were not alignable between the outgroup and ingroup taxa and were thus treated as missing data for the outgroup. Indels associated with unalignable regions in the ITS1 and ITS2 genes in the ingroup taxa were treated as missing data in maximum likelihood datasets, transformed into either binary or multistate traits in the Bayesian inference (Table 3).

Data blocks were set up according to the rDNA gene (18s, ITS1, 5.8s, ITS2, 28s), rDNA indels (included in a single block), or ND1 codon position (1^{st}, 2^{nd}, 3^{rd}). To correctly annotate the rDNA regions, I compared the ribosomal sequence to the complete nuclear ribosomal RNA sequences of 2 species of the digenean *Diplostomum* (Brabec et al. 2015). I then used PartitionFinder v.1.1.0 (Lanfear et al. 2012) to determine the best partition scheme (according to BIC) and model of evolution for each partition. I used the following parameters in ParitionFinder searches: branch lengths = linked; models = mrbayes or raxml (run individually); model selection = BIC; data blocks defined by gene (rDNA) and codon position (ND1 mtDNA); and search = all. The most appropriate model of DNA substitution was determined in PartitionFinder, the partitioning schemes of the individual rDNA and ND1 datasets were subsequently applied to the applicable regions of the concatenated dataset.

I conducted maximum likelihood (ML) searches in RAxML v.8 (Stamatakis 2014): analyses were run under the Rapid Bootstrap Algorithm and consisted of 1,000 bootstrap replicates followed by a thorough ML search under the GTR + Γ model as determined by PartitionFinder (bootstrap values ≥ 70 indicating strong support [Hillis

Table 3. Insertion/deletion events associated with unalignable regions in the ITS1 and ITS2 rDNA regions of *Alloglossidium* species. The size is given in base pairs (bp) and the position refers to the position on the aligned concatenated data matrix. Treatment of these events (transformed into multistate or binary data; position of alignment after transformation) is also provided.

Data Matrix Position	Size (bp)	Description	Matrix Position after Transformation
ITS1			
904-912	6	Coded as multistate with 10 states (0-9)	4164
927-929	2	Insertion coded as binary (0 or 1)	4165
973-986	14	No insertion (coded as 0) 12 bp insertion in *A. anomophagis* (coded as 1) 14 bp insertion in *A. progeneticum*, *A. renale,* and A. *greeri* (coded as 2)	4166
997-999	3	Coded as multistate with 8 states (0-7)	4167
1032-1035	4	Coded as multistate with 6 states (0-5)	4168
1386-1390	5	Coded as multistate with 4 states (0-3)	4169
1421-1422	2	Insertion coded as binary (0 or 1)	4170
ITS2			
1999	1	Insertion coded as binary (0 or 1)	4171
2008	1	Insertion coded as binary (0 or 1)	4172
2099	1	Insertion coded as binary (0 or 1)	4173
2197	1	Deletion coded as binary (0 or 1)	4174
2659-2660	2	Coded as multistate with 4 states (0-3)	4175

and Bull, 1993]). I conducted Bayesian phylogenetic reconstructions with the MrBayes v3.2 (Huelsenbeck and Ronquist 2001) package. So that each partition was allocated its own set of parameters, statefreq (state frequency), tratio (t-ratio), revmat (transition matrices), pinvar (proportion of invariant sites), and/or shape (gamma shapes) were coded as unlinked and the ratepr parameter was set to variable to allow for the evolution of partitions under different rates. In the concatenated analysis, I allowed for the estimation of separate branch lengths (brlens) between the rDNA or mtDNA partitions. PartionFinder identified 3 partitions for MrBayes: 1) 18s and 5.8s (K2P; Kimura, 1980), 2) ITS1 and ITS2 (SYM + Γ; the GTR model with stationary state frequencies fixed to be equal), and 3) 28s (HKY = Γ; Hasegawa et al. 1985). For the mtDNA dataset PartitionFinder identified 2 partitions: 1) 1st and 2nd codons combined (HKY+ Γ + I) and 2) 3rd codon alone (HKY+ Γ). The multistate/binary traits included as a sixth partition and were estimated using Γ rate only. Two parallel runs each with 4 chains were run over 4 million generations, sampled every 1000 steps. The run length was deemed appropriate based on the convergence of the parallel runs (standard deviation of split frequencies < 0.01). Also, for all model parameters, the potential scale reduction factors approximated 1, thus indicating a good sample from the posterior probability distribution. After reaching stability, I discarded the first 25% of trees as burn-in and calculated a 50% majority-rule consensus tree (posterior probabilities ≥ 95% indicating strong support).

Given the high support for the clade containing *A. progeneticum*, *A. renale*, and *A. greeri*, coupled with recent surveys results (Kasl et al. 2015) that highlighted the

intraspecific variation in life cycle pattern of *A. progeneticum*, a further analysis was run to assess the evolution of precociousness within a single clade. The ND1 mtDNA sequences of all known *A. progeneticum* haplotypes (see Chapter IV; GenBank accession nos. KT455707-KT45582, Kasl et al. 2015; KU728083-KU728090, McAllister et al. *in press*) were compiled with specimens of *A. progeneticum, A. renale, A. greeri*, and *A. geminum* (chosen as the outgroup due to its position as ancestral to *Alloglossidium* species in fish and crustacean hosts). A 676 bp dataset was created and the maximum likelihood (ML) and Bayesian reconstructions were analyzed using the same parameters specified for the ND1-only dataset above.

Reconstructing the Evolution of Life Cycle Complexity

Character states associated with life cycle complexity and coded for use in ancestral state reconstructions were (0) obligate 2-host crustacean, (1) facultative precocious, (2) obligate 2-host leech, and (3) obligate 3-host (Table 1). The life history states of the outgroup were coded as missing data to prevent artificially influencing the evolution of traits within the ingroup. Ancestral states reconstruction was carried out using the maximum likelihood method, with transitions between each character state coded as equally likely, in Mesquite v3.04 (Maddison and Maddison 2010). The consensus tree produced from the Bayesian analysis (using PartionFinder delineated partitions) of the concatenated dataset was read into Mesquite and character history was plotted using the trace character history command. As the trace character history command approximates the likelihood of the ancestral state using branch length, I ran the reconstruction using branch length estimates from both the rDNA and

ND1independent (i.e., unlinked) assessments. A probability ≥ 0.60 for a character is typically considered to be the most probable state at that node (Feutry et al. 2013; Chauhan and Pandey 2014).

Results

Molecular Variation

Among *Alloglossidium* sequences, 2182 nucleotides were originally aligned in the ribosomal DNA dataset. There were 42 sites that could not be unambiguously aligned due to associated indels and thus, were recoded as either binary or multistate traits for analyses (Table 3). Within the 2140 remaining bases, after excluding the above mentioned regions, 274 sites were found to be polymorphic with 28 singleton and 246 informative polymorphisms. Twenty-four haplotypes were found from 49 individuals. Among these 24 unique haplotypes, there were the 17 nominal species included in this study and 2 new species that have yet to be described (*Alloglossidium n. sp.* 1 and *Alloglossidium n. sp.* 2) (Fig. S1). With the exception of 3 morphologically identified species, all ribosomal sequences were identical to other specimens sharing the same morphological ID. Among the two specimens identified as *A. richardsoni*, there was a single A/G substitution in the ITS1 region. There was an *A. fonti* clade with 3 haplotypes split into two groups: 2 haplotypes with a northern range (Vermont/Wisconsin; labels FO01 and FO02 in Fig. S1) and 1 haplotype with a southern range (Texas/Louisiana; labels FO03 and FO04 in Fig. S1). Specimens morphologically identified as *A. kenti* were found to represent 3 highly diverged haplotypes (Fig. S1). Specimens sampled from the type locality of *A. kenti* (Money, Mississippi; labels MS01-

03 in Fig. S1) did not group as sister to the specimens retrieved from Texas, Arkansas, or New York (labels TK01-05 in Fig. S1). These latter specimens had rDNA sequence similar to that reported in Tkach and Mills (2011) (Chapter III). Additionally, the *A. kenti* from New York (labels TK04-05 in Fig. S1) represented a different haplotype (2 nucleotide difference) than was found in Arkansas and Texas (labels TK01-03 in Fig. S1). Henceforth, I denote the *A. kenti* lineages as *A. kenti* MS for those obtained from the type locality and *A. kenti* TK for those whose rDNA corresponds to the sequence used by Tkach and Mills (2011) to resurrect the species. The variation associated with *A. fonti* and *A. kenti* are discussed in further detail in Chapter III.

The ND1 mtDNA region consisted of 676 bp, contained no gaps in either ingroup or outgroup species, and yielded 40 unique haplotypes among the ingroup specimens. Within both the ingroup and outgroup, no premature stop codons were found after translation of the sequences (using amino acid translation code 9 on GenBank). Table S1 reports pairwise ND1 *p*-distances. For specimens that I initially identified to species based on morphology, I found the following distances. In the case of the *A. fonti* specimens, three individuals (FO01, Vermont; FO02, Wisconsin; and FO04, Texas) have *p*-distances between 2.2 and 3.3% when compared against each other, but pairwise comparisons with the fourth individual (FO03, Louisiana) yields *p*-distances of 5.5-5.9%. Additionally, among the individuals associated with the *A. kenti* TK designation, the 3 individuals from Texas (TK01, TK02) and Arkansas (TK03) differed by <1%, whereas the 2 individuals form New York (TK04 and TK05) showed differences of 6.8-7.7% to TK01-TK03. This latter result suggests that there may be unrecognized cryptic

speciation within this lineage (for further discussion see Chapter III). Interspecific comparisons of the ND1 typically showed a *p*-distance between 9.6-21.4%, (Table S1). However, comparisons of *A. hamrumi*, *A. hirudicola*, and *A. turnbulli* were found to have *p*-differences between 1.6-4.9%, ranging as low as 1.6-2.4% between *A. hamrumi* and the 2 *A. turnbulli* specimens (Table S1). This latter result is in stark contrast to the large rDNA divergence found for *A. turnbulli* specimens (Fig. S1)

After accounting for identical sequences, the 49 original *Alloglossidium* individuals were condensed down to 40 unique sequences, yielding a concatenated rDNA-ND1alignment of 2820 bp, after removing gaps.

Phylogenetic Relationships

Although phylogenetic reconstructions did not qualitatively differ between analytical methods (i.e., the topologies produced by MrBayes and RAxML were similar, with minor variations in node support values), I did find incongruences between the individual rDNA and ND1 data sets. In particular, the rDNA 50% majority consensus tree had greater nodal support values, especially with regards to deep divergences. Both MrBayes and RAxML rDNA trees had support for all species using leeches as a final host in a single clade (Clade B; PP: 92; BS: 95) that diverged from all remaining *Alloglossidium* species (using crustaceans and ictalurids as a definitive host; Clade A; PP: 100; BS: 85; Fig. S1). The support for the monophyly of all currently recognized species of *Alloglossidium* included in this study was high (>98% for both analyses). Basal to the species in Clade A was *A. geminum* (PP: 100; BS: 85) with a clade containing *A. kenti* MS and *Alloglossidium n. sp. 1* (PP: 100; BS: 98) as sister to the

remaining 9 species. With the exception of 3 clades, the stability of the remaining interspecific internal nodes is dependent on analyses method, with RAxML support values often leading to collapsed branchs (PP: 60-85; BS: <50-83). However, the rDNA data do support a designation of *A. progeneticum*, *A. renale*, and *A. greeri* as clade (Clade II; PP: 100; BS: 100), *A. cardicolum* and *A. dolandi* as a clade (Clade I; PP: 100; BS: 100), and *A. anomophagis* and *A. fonti* as sister taxa (PP: 97; BS: 87).

The ND1 mtDNA data set does not provide resolution for deeper divergences as that of the rDNA. In fact, the support values were so low as to create a polytomy at the base of the tree (Fig. S2). There are some well-supported internal nodes that are also shared with the rDNA. All of the species using leech final hosts group together as a clade (Clade B; PP: 74; BS: 86), though there are some differences in the relationships among the most recently derived leech species between the 2 genetic marker datasets. In particular, *A. turnbulli* has low mitochondrial genetic divergence to some other species that uses leeches compared to its large rDNA divergence (compare Figs. S1 and S2). Additionally, *A. cardicolum* and *A. dolandi* are still supported as sister taxa (Clade I; PP: 99; BS: 82) and the Clade II grouping of *A. progenenticum*, *A. renale*, and *A. greeri* has strong support (PP: 100; BS 99). Most problematic, however, is the placement of *A. anomophagis* as sister to *A. floridense* (PP: 98; BS: 61), which could be due long-branch attraction (rapidly evolving mtDNA lineages) associated with these species (Bergsten 2005).

The relationships found in the concatenated tree more closely adhere to those found in the rDNA where there is strong support for monophyly of genus (>98% for

both analyses) and for an early divergence between those species in leeches (Clade B; PP: 99; BS: 86) and the remaining *Alloglossidium* species (Clade A; PP: 100; BS: 85; Fig. 4). Within Clade A, *A. geminum* was again placed as the basal species. The next most basal group within Clade A contains the sister taxa *A. kenti* MS and *Alloglossidium n. sp. 1* (PP: 100; BS: 100). The concatenated data set does provide support to for *A. kenti* TK to fall basal to the remaining species (PP: 100; BS: 84). Additionally, there is support for a clade delineating *A. anomophagis* and *A. fonti* as sister taxa (PP: 98; BS: 71), *A. cardicolum* and *A. dolandi* as sister taxa (Clade I; PP: 100; BS: 100), and the "*progeneticum* clade" of *A. progeneticum*, *A. renale*, and *A. greeri* (Clade II; PP: 100; BS: 100).

Within the Clade II-only ND1 analyses of *A. progeneticum*, *A. renale,* and *A. greeri*, the precocious species (i.e. the facultative *A. progeneticum* haplotypes associated with distinct river drainages, *A. greeri*, and *A. renale*) were found to form a polytomy (Fig. 5). The ancestral condition was found to be the obligate 3-host pattern of the non-precocious *A. progeneticum* populations. *Alloglossidium renale* and *A. greeri* were found to fall out in a highly supported clade (PP: 100; BS: 99) and the 2 haplotypes from the Oconee River drainage in Georgia were also formed a highly supported clade (PP: 100; BS: 96). For further information on the associations within the *A. progeneticum* samples, see Chapter IV.

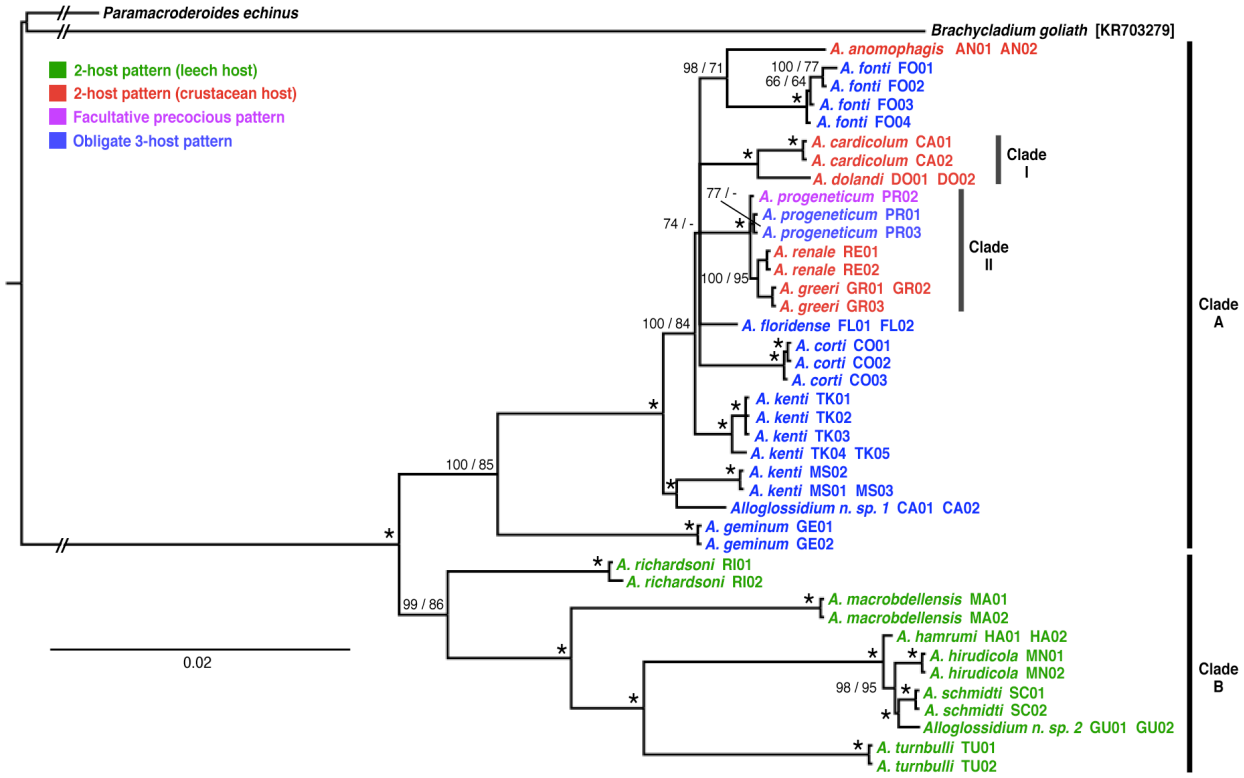

Figure 3. The 50% majority rule consensus tree obtained from Bayesian analysis of the concatenated ribosomal (18s-ITS1-5.8s-ITS2-28s rDNA) and mitochondrial (ND1 mtDNA) dataset. Numbers at nodes correspond to the Bayesian posterior probability / RAxML maximum likelihood bootstrap support percentages; * denote nodal support values > 98% for both analyses, - denotes ML nodal support values of < 50%. The outgroup taxa (*P. echinus* and *B. goliath*) have been collapsed and branch length shortened. Colors denote life cycle pattern (green = 2-host life cycle maturing in leech hosts; red = 2-host life cycle using crustacean final hosts; purple = facultative precocious; blue = obligate 3-host). Clades A, B, I, and II are explained further in the text.

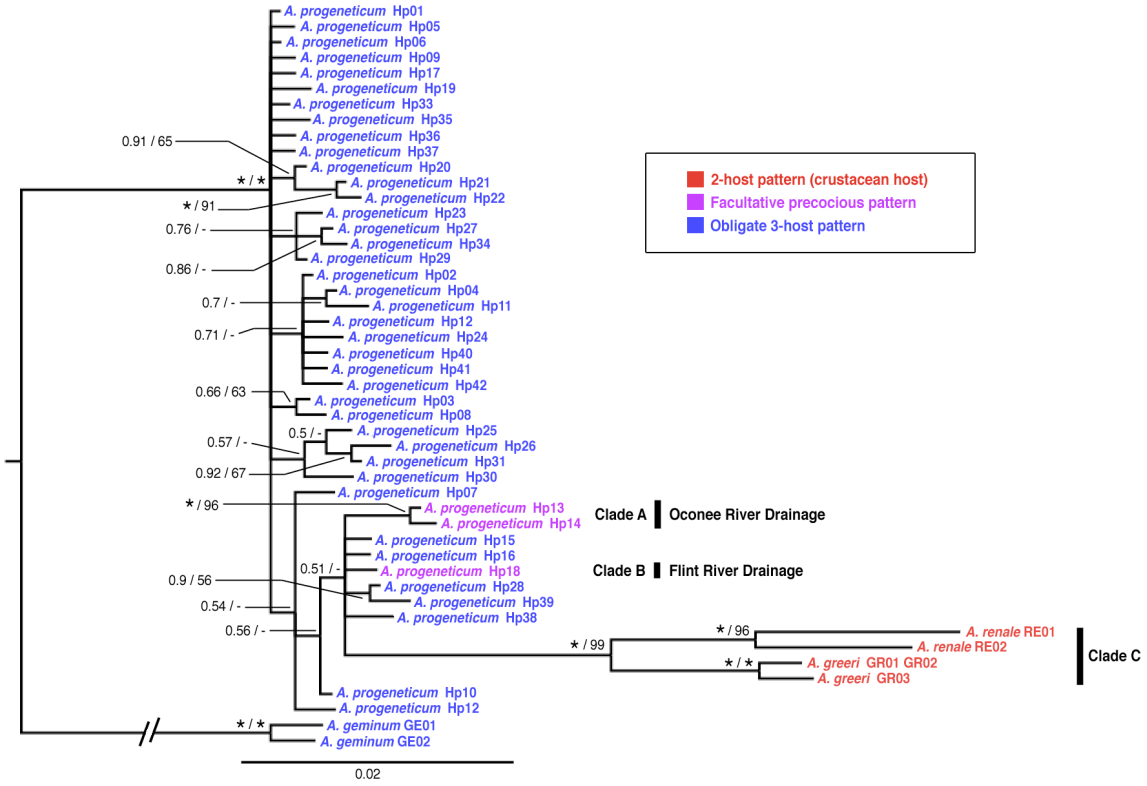

Figure 4. The 50% majority rule consensus tree obtained from the Bayesian analysis of mitochondrial (ND1 mtDNA) haplotypes associated with *A. progeneticum*, *A. renale*, and *A. greeri*. Numbers at nodes correspond to the Bayesian posterior probability / RAxML maximum likelihood (ML) bootstrap support percentages; * denote nodal support values of 100%, (-) denotes ML nodal support values of < 50%. The outgroup taxa (*A. geminum*; representing an ancestral, obligate 3-host life cycle) has been collapsed and branch length shortened. Colors denote life cycle pattern (red = 2-host life cycle using crustacean final hosts; purple = facultative precocious; blue = obligate 3-host). Clade A (Oconee River), B (Flint River), and C (loss of encysted metacercarial) are discussed further in the text.

Inference of Ancestral States

Given that the mtDNA had poor resolution for deeper relationships, I relied on the concatenated data set to draw most of the inferences about life cycle evolution. Here I present the visual phylogram of the host-number analyses with discussion of the statistical results of each analysis (incorporating the individual branch lengths associated with the rDNA and ND1 associated branch length estimations) at the nodes corresponding to evolutionary events (Fig. 6). The ancestral state at the split between the major *Alloglossidium* clades (Node A) was found to be ambiguous: obligate 3-host had 61% using mtDNA branch lengths or 32% with rDNA branch lengths whereas the obligate 2-host leech had 36% using mtDNA branch lengths or 59% with rDNA branch lengths; Fig. 6). Node F was found to have higher ML predicted probabilities relating to the 2-host, leech pattern: the 2-host, leech pattern, was supported using both the mtDNA or rDNA branch lengths (83% and 89%, respectively) whereas the obligate 3-host pattern had much lower support (mtDNA = 12%; rDNA = 9%; Fig. 6). The remaining labeled nodes had the obligate 3-host life cycle as the most probable ancestral states (B: 3-host = 0.79-0.96; C: 3-host = 0.99; D: 3-host = 0.99; and E: 3-host = 0.94); the independent loss of the 3rd host is therefore inferred at nodes C, D, and E (Fig. 6).

Discussion

Interrelationships of Alloglossidium spp.

A molecular phylogeny of the genus *Alloglossidium* is a necessary first step in asking questions related to the evolution of changes in life cycle complexity. Without wading into the debate on the superiority of molecular (Scotland et al. 2003) or

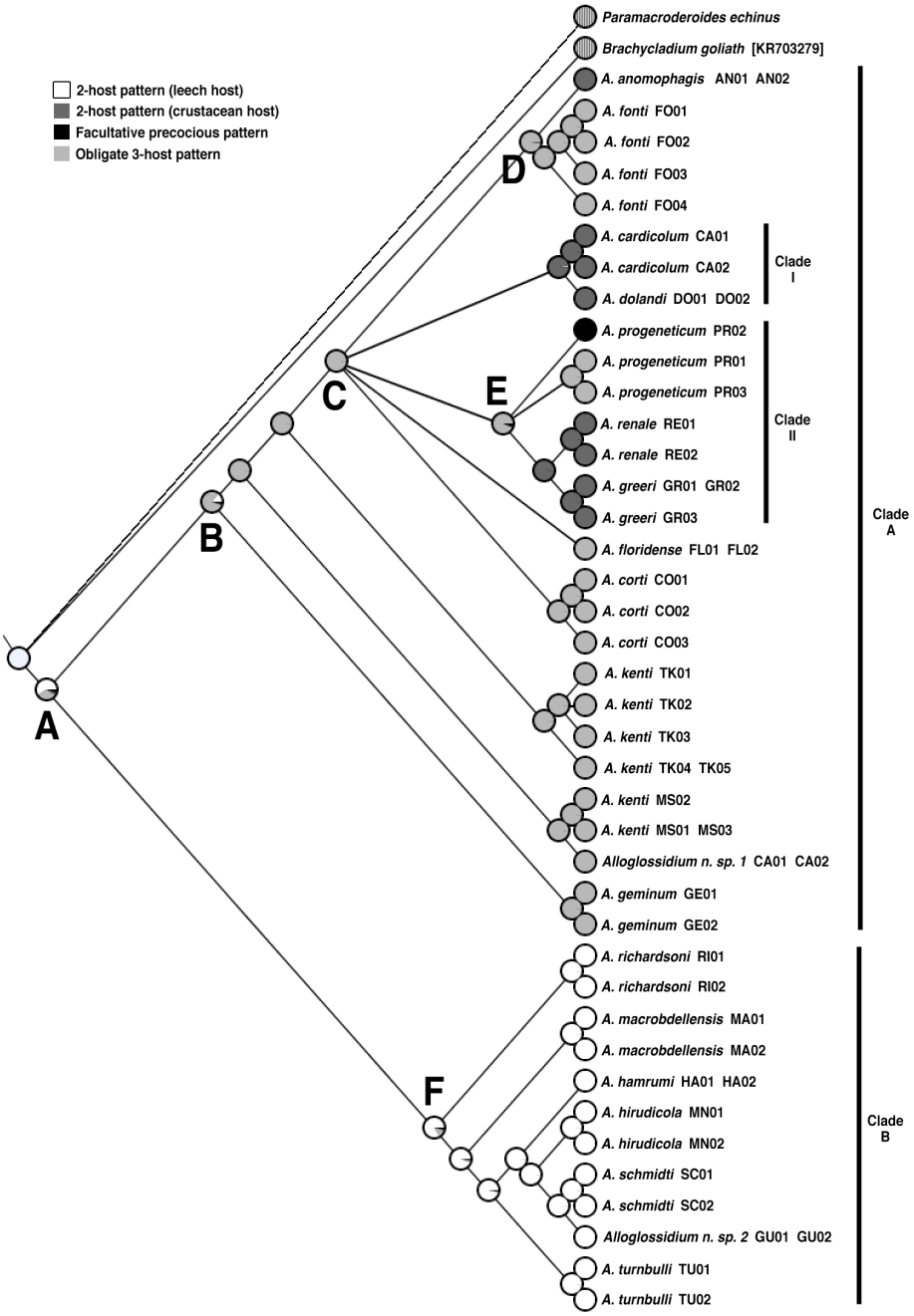

Figure 5. Reconstructed evolutionary history of changes in life cycle complexity [4 states: 2-host (leech host), 2-host (crustacean host), facultative precocious, obligate 3 host] using the maximum likelihood (ML) method implemented in Mesquite v3.04 and plotted on the 50% majority concensus tree from the concatenated Bayesian analysis using rDNA branch length estimates (see Fig. 4). The highest ML scores at each of the indicated nodes (A-F) are discussed in text**.**

morphological data (rebuttal, Jenner 2004) in creating the most valid phylogenetic reconstructions, the molecular phylogeny provides one major advance over the previously proposed morphological and life history-based phylogenies of *Alloglossidium*: a means of creating an independent assessment of relationships that can subsequently be used to address character trait evolution. Indeed, the incorporation of 2 independently evolving genes (ribosomal vs. mitochondrial) helps to highlight the most significant clades and, when combined, allows for better resolution across the phylogeny.

There are some similarities between the molecular phylogenies presented in this study and the morphological phylogeny presented by Smythe and Font (2001). First, the molecular phylogeny does support Smythe and Font's (2001) placement of *Alloglosoides dolandi*, *Alloglossoides cardicola*, and *Hirudicolotrema richardsoni* within the genus *Alloglossidium*. *Alloglossidium anomophagis* was not included in the analysis of Smythe and Font (2001). Herein, my molecular phylogeny supports the inclusion of this species within the genus. Second, those species using a 2-host pattern with leeches as the definitive host form a monophyletic clade (Clade B). Third, all phylogenies support *A. cardicolum* and *A. dolandi* as sister taxa (Clade I; Figs. 3 and 4).

The molecular phylogeny, however, also had major differences compared to the morphological and life-history based phylogeny of Smythe and Font (2001). First, the concatenated dataset supported an early divergence between species in leeches (Clade B) and all other species. *Alloglossidium geminum* was found to be basal in Clade A, which contains all of the non-leech associated *Alloglossidium* species. (Fig. 4). In contrast,

Smythe and Font's (2001) analysis places *A. geminum* ancestral to all *Alloglossidium* species with those species from leeches forming the most recently diverged clade (Fig. 3). Furthermore, Smythe and Font (2001) placed *A. richardsoni* as the most recently derived species, whereas the molecular data suggest it is actually the most basal member of the leech clade (Fig. 3; Clade B, Fig. 4). Finally, sequence data support the delineation of a '*progeneticum*' clade (Clade II; comprised of *A. progeneticum*, *A. renale*, and *A. greeri* as sister taxa). Smythe and Font's (2001) analysis does not show *A. progeneticum* to be an exclusive sister taxa of *A. renale* and *A. greeri*. Rather their analysis shows a successive splitting of *A. renale* and *A. greeri* of the main branch.

Reconstructing the Evolution of Life Cycle Complexity

When using the morphology and life history-based tree of Smythe and Font (2001), the most parsimonious evolution of life cycle complexity is as follows: 1) the obligate 3-host pattern (catfish definitive host) is ancestral, 2) the loss of the fish host occurred according to a progressive event from a facultative precocious to an obligate 2-host pattern with a crustacean definitive host, and 3) a host-switching event from crustacean to leech final hosts. The latter event would have had to entail the regaining of the cyst development in species using leeches as *A. renale*, *A. greeri*, *A. cardicolum* and *A. dolandi* do not have an encysted stage. Ancestral state reconstructions based on the molecular phylogeny do not agree with any of these hypotheses. First, the analyses showed a well-supported ancient divergence between the clades belonging to leeches (Clade B) and all other *Alloglossidium* species (Clade A). The probability of the ancestral state at the node where these clades split, however, is equivocal as to whether

the obligate 3-host or 2-host leech pattern is the ancestral state. Thus, further analyses (e.g., using a more closely related outgroup species might give more confidence in what the ancestral trait was to the genus itself) would be needed to tackle Riggs and Ulmer's (1983) hypothesis that leeches served as the ancestral host. However, based on the 2 major clade designations, I can say with certainty that the ancestral condition of Clade A is an obligate 3-host (ML > 0.99), fish final host (ML > 0.99) pattern, while the ancestral, and indeed only life cycle, of Clade B is the 2-host, leech final host condition (ML > 0.84). Furthermore, these analyses provide support for multiple losses of the fish definitive host throughout Clade A: ML probabilities of the nodes immediately preceding the delineation of precocious species (Nodes D: *A. anomophagis*; Node C: *A. cardicolum* and *A. dolandi*; and Node E: *A. progeneticum*, *A. renale,* and *A. greeri*; Fig. 6) are always predicted to have the obligate 3-host life cycle as the ancestral trait (ML > 0.95). Thus, given the most recent common ancestor of each precocious life cycle is the obligate 3-host pattern, then I can infer that there are 3 independent origins of precociousness. This results in 3 complete losses of the ictalurid host and 1 facultative use of an ictalurid host. I note that Font (1980) hypothesized an evolutionary progression from obligate 3-host species to a facultative precocious life cycle (*A. progeneticum*) to an obligate 2-host life cycle (*A. renale*) maturing in crustaceans. At the time, however, *A. renale*, *A. greeri*, *A. cardicolum* and *A. dolandi* were not yet described. The concatenated phylogeny shows the species with a facultative life cycle (*A. progeneticum*) and obligate 2-host life cycle (*A. renale* and *A. greeri*) are closely related. However, the ND1-only phylogeny (Fig. 5) incorporating haplotypes across

populations of *A. progeneticum* with different life histories shows possible multiple origins of facultative and obligate 2-host life cycles emerging from the ancestral obligate 3-host life cycle of the '*progeneticum'* clade (Clade II) (see also Chapter IV). Moreover, *A. cardicolum* and *A. dolandi* form a separate clade. Thus, a single stepwise-progression in a the loss of the vertebrate host is not supported.

Cryptic Species

The survey work coupled with the molecular phylogeny of the nominal species *Alloglossidium* also highlights potential cryptic species. This is particularly relevant for the species reported from catfishes, which have subtle morphological variation and have existed in a taxonomic limbo (due to unrecognized cryptic species) for the better part of 80 years (Tkach and Mills 2011; Kasl et al. 2014; Kasl et al. 2015). As suggested by Vilas et al. (2005), a level of 5% *p*-distance in mtDNA (ND1 or CO1) should raise a flag with regards to the possible presence of cryptic species. I found evidence for cryptic diversity in 2 lineages. In the case of *A. fonti*, ND1 *p*-distances of 5.5%-5.9% separating the individual from Louisiana from those sampled elsewhere (Vermont, Wisconsin, and Texas) are at the threshold for intraspecific variation. However, the Louisiana and Texas sample share the same rDNA haplotype. This could be case of incomplete lineage sorting among distinct species. More samples are required to determine if cryptic lineages are present in the *A. fonti* clade. Samples designated as *A. kenti* present a more complex situation. Ribosomal sequence clearly distinguishes the *A. kenti* sampled from the type locality of *A. kenti* (MS) from those specimens that were resurrected as *A. kenti* (TK) by Tkach and Mills (2011). Furthermore, there appears to be another level of

cryptic variation within the lineage associated with the *A. kenti* TK samples; specimens from New York were found to have ND1 divergence of 6.8-7.7% when compared to other *A. kenti* TK samples (Table S1). This mitochondrial divergence is supported by different rDNA haplotypes between the New York samples and other *A. kenti* TK samples. This cryptic diversity and the need for the redescription/description of *Alloglossidium* species from fishes is further addressed in Chapter III where diagnostic morphological traits are addressed. Finally, the low *p*-distances among ND1 mtDNA sequences of some *Alloglossidium* species from leeches when compared to the rDNA divergence patterns is also notable, especially with regards to *A. turnbilli*. Further work is needed to elucidate the causes and validity of those relationships due to the divergence found in the rDNA regions of the same species. However, I do note that some of the species that use leeches can be found concurrently in the same host individual. Thus, historical introgression events remain a possibility to explain the discordant mtDNA and rDNA patterns.

Caveats and Summary

Recognizing the issues involving incongruences among these rDNA and mtDNA trees, I nevertheless note that the rDNA and concatenated molecular phylogenies show strong support for an early evolutionary split between those species that utilize leech final hosts and those that parasitize ictalurids and crayfish (Fig. 4; Fig. S1). The ND1 mtDNA is not as informative, instead forming a polytomy at the base of the tree (Fig. S2). This issue may be resolved with a more closely related outgroup or the addition of more informative loci.

Given the widely recognized potential for the genus *Alloglossidium* to serve as a model system for life cycle evolution (Brooks 2003), this study presents a much-needed framework with which to address the causes and consequences of changes in life-cycle complexity. The updated *Alloglossidium* phylogeny not only provides the most current and complete analysis of species relationships (including 17 of the 18 nominal taxa), but also restructures our knowledge of those relationships within *Alloglossidium*. The illumination of multiple origins of precociousness also provokes more questions: What is the relationship between morphology and life cycle pattern? For example, does musculature or spination change depending on location within host? How does reaching sexual maturation when still encysted (i.e., forced self-mating) effect sex allocation? Ultimately, this study establishes a phylogenetic framework with which to address future studies in character evolution while also reemphasizing the potential for the genus *Alloglossidium* to be used as a model to assess the ecological and evolutionary impacts of complex life cycles.

CHAPTER III

ELUCIDATING CRYPTIC DIVERSITY IN *ALLOGLOSSIDIUM* SPP. FROM FISHES: THE TAXONOMIC IMPORTANCE OF CONNECTING MOLECULAR IDENTIFICATIONS TO MORPHOLOGICAL VARIATION

Introduction

The discovery of cryptic species (i.e. morphologically similar, but genetically distinct taxa) has a direct impact on our assessment of parasite biodiversity (Poulin and Morand 2004). Without the ability to definitively recognize a species as unique, we run the risk of underestimating parasite species richness. The issue of unrecognized diversity is compounded due to the limited number of morphological traits found in parasites. As such, it is not always known if the observed phenotypic variation is taxonomically valid (i.e., diagnostically representative of distinct species). Morphological differences may be due to environmental factors (e.g., host species, density-dependence), investigator induced (e.g., fixation and mounting methods; Criscione and Font 2001), or even represent underlying intraspecific genetic variation (Perkins et al. 2011). With such subtle variation often found among parasites, distinct species are likely to go unrecognized when using morphology alone as a means of delimitation. Therefore, it is not surprising that the finding of cryptic species is an increasingly well-recognized taxonomic issue among digenean parasites (Criscione and Blouin 2004; Locke et al. 2010; Pérez-Ponce de León and Nadler 2010; Nadler and Pérez-Ponce de León 2011; Poulin 2011). As molecular survey, population genetic, and

phylogeographical studies increase, it is likely that more cryptic species will be encountered (Criscione et al 2005). Research on cryptic species thus necessitates a 3 step approach: 1) recognition of potential cryptic species, 2) delimitation through hypothesis testing (i.e., using a null hypothesis of a single species), and 3) formal description and naming of the new taxa (Pérez-Ponce de León and Nadler 2010). Recent surveys and phylogenetic work in the digenean genus *Alloglossidium* Simer, 1929 (Chapter II) has revealed cryptic diversity that necessitates taxonomic revision.

To highlight current taxonomic problems, I herein provide a brief history of recent molecular work within the genus *Alloglossidium*. Prior to Tkach & Mills (2011) (the first study using a combined molecular and morphological taxonomic approach in the genus *Alloglossidium*), only 3 species were recognized from catfish hosts: *Alloglossidium corti* (Lamont, 1921), *Alloglossidium geminum* (Mueller, 1930), and *Alloglossidium progeneticum* (Font and Corkum 1975). Tkach and Mills' (2011) application of ribosomal sequence data was instrumental in not only delimiting species, but also highlighted the necessity of using the genetic data to corroborate observed morphological differences which may otherwise be wrongly attributed as intraspecific variation. As a result of their study, a fourth species was described (*Alloglossidium fonti*) and a fifth species was resurrected (*Alloglossidium kenti*). The case of *A. kenti*, was particularly interesting, as the species had long been synonymized under the *A. corti* designation (Van Cleave and Mueller 1934); Tkach and Mills even admitted that had there not been extensive sequence variation in their *A. kenti* specimens, it is likely that they would not have recognized the individuals as a different species. However, it is

important to note that Tkach and Mills did not obtain *A. kenti* specimens from the type locality, though they did sample in the surrounding geographic regions.

Kasl et al. (2014) were the next to apply a combined morphological and molecular approach to species delineation, describing *Alloglossidium floridense*. In doing so, they highlighted that that complete morphological information is often missing from the literature. This is due to the fact that original descriptions were sometimes based off of unsuitable preparations of material (*A. corti*, as discussed by Mueller 1930) or lacked certain details (i.e., the descriptions of *A. corti*, *A. geminum*, and *A. kenti* from the 20s and 30s). In addition, Kasl et al. (2014) noted many prior identifications fall under the umbrella of *A. corti*. Thus, current records of the geographic distributions, hosts, and morphological variation associated with *A. corti* are likely to contain errors due to the unrecognized cryptic diversity that has come to light with molecular data. For instance, *Alloglossidium progeneticum* was recently found to have unrecognized intraspecific morphological and life history variation across the southeastern U.S. (Kasl et al. 2015); it is likely that some of the sites previously identified as *A. corti* may in fact be *A. progeneticum*. Finally, recent survey work and a molecular phylogenetic reassessment of the genus (Chapter II) indicated new cryptic lineages, including specimens from the type locality of *A. kenti* that were genetically divergent from the sequences provided by Tkach and Mills (2011).

The genus has been advocated as a potential model for life cycle evolution (Brooks 2003; Chapter2). However, the usefulness of the *Alloglossidium* system hinges on establishing the correct taxonomy. Though this chapter focuses on the recognition of

the cryptic morphology of *Alloglossidium* species from fishes (and the determination of what morphological characters, if any, are diagnostic), it does not serve as a valid species description and the information conveyed herein should not be treated as such. However, I note that this chapter serves as a precursor to that necessary taxonomic revision. In particular, I address the case of "*A. kenti*", which includes two molecularly distinct lineages, associated with 1) the type locality and 2) the specimens reported in Tkach and Mills' (2011). In addition, I reexamine *A. corti* and *A. geminum*, as prior literature on these species may have included trait variation of other species and/or have species descriptions that are incomplete. Lastly, I discuss a new lineage from Canada that was identified via phylogenetic assessment (Chapter II). In total, these data suggest a total of 8 species from catfishes.

Methods

Sampling of *Alloglossidium* species was conducted as part of a large-scale biodiversity survey of the eastern two-thirds of the United States and southern Canada (see Chapter II for additional locality information). All specimens were initially identified using morphological descriptions, killed with hot water or heat-fixed underneath a cover slip without pressure to ventrodorsally orient the worms, and stored in either 70% or 95% ethanol for potential downstream molecular analyses. Subsequently, specimens used for morphometric analyses were stained using Semichon's acetic carmine, dehydrated in a graded ethanol series, cleared using xylene, and mounted on slides using Damar gum. Figures were made with the aid of a drawing tube. All measurements were taken from mounted specimens and are given in

micrometers; the mean is given for each measurement followed by the range in parentheses.

Based on prior phylogenetic work (Chapter II), 2 new species of *Alloglossidium* from ictalurids were identified. I therefore refer to these new species as *Alloglossidium n. sp.* A and *Alloglossidium n. sp*. B, and state that formal species descriptions will be submitted in a future manuscript. *Alloglossidium n. sp.* A, corresponds molecularly and morphologically to the specimens that Tkach and Mills (2011) used to resurrect *A. kenti*; I obtained specimens deposited by Tkach and Mills to provide measurements of *Alloglossidium n. sp.* A and also included specimens from channel catfish, *Ictalurus punctatus* (Rafinesque, 1818), from Little Brazos River (TX; 30° 38.485 N 96° 31.222333' W) that had the same rDNA sequence as those given by Tkach and Mills. Thus, specimens from this location were sampled and included in the comparative morphometric analyses and the ultimate species description. Specimens of *Alloglossidium n. sp.* B were recovered from the intestines of channel catfish retrieved from the St. Lawrence River, Quebec, Canada (45° 18.96' N 73° 52.74' W).

Molecular Analyses

Ribosomal sequences were used from specimens of *A. corti* and *A. geminum* that had been morphological identified to species (see Chapter II). These sequences were chosen to represent locations associated with the morphological analyses. Four specimens each of *A. kenti* (from the type locality in Money, MS), *Alloglossidium n. sp.* A, and *Alloglossidium n. sp.* B were used to obtain rDNA sequence data for comparative molecular analysis. I also compared the sequences from my survey work to those

previously deposited in GenBank (Tkach and Mills 2011: JF440783.1, *A. corti*; JF440765.1and JF44767.1, from species description, *A. fonti*; JF440771.1, *A. geminum*; JF440808.1, *A. kenti*; Kasl et al. 2014: KC812276.1, from species description, *A. floridense*) for further delimitation. However, I note that Tkach and Mills (2011) did not sequence the ND1 mitochondrial region, so only the ribosomal sequence was used in comparisons. DNA was extracted by placing individual worms in 200 µl of 5% chelex containing 0.2 mg/ml of proteinase K, incubated for 2 hr at 56 C and boiled at 100 C for 8 min. Polymerase chain reaction (PCR) amplifications were performed resulting in DNA fragments of approximately 2,450 base pairs spanning from the 3' end of the 18s nuclear rDNA gene through the internal transcribed spacer region 1 (ITS1), 5.8s, and ITS2 genes to the 5' end of the 28s gene (including domains D1-D3). Twenty-five µl reactions containing 3 µl of template extract, 15.25 µl water, 2.5 µl 10x buffer, 2.5 µl $MgCl_2$ [25 mM], 0.5 µl dNTP [10 mM/each], 0.5 µl of each primer [20 µM], and 0.25 µl of OmegaTaq polymerase [5 units/µl] were used following the thermocycler profile described in Tkach et al. (2003). An *Alloglossidium*-based primer near the 3' end of the 18s, CC41 (5'-GATTGAATGGTTTAGCAAGG-3') (Kasl et al. 2014), and the general trematode 28s reverse primer 1,500R (5'-GCTATCCTGAGGGAAACTTCG-3') of Olson et al. (2003) were used for the PCR. In addition to the PCR primers, 3 internal forward primers [CC48 (5'-TGTCGATGAAGAGTGCAGC-3') from Kasl et al. (2014), and 900F (5'-CCGTCTTGAAACACGGACCAAG-3') and 300F (5'-CAAGTACCGTGAGGGAAAGTT G-3') from Olson et al. (2003)] were used for sequencing. PCR products were purified using the Ultra Clean PCR clean-up Kit (MO

BIO Laboratories, Inc., Solana Beach, California) and sent to the DNA Analysis Facility on Science Hill at Yale University (New Haven, Connecticut) for sequencing. Contiguous sequences from individuals were assembled using Sequencher™ (GeneCodes Corp., ver. 4.1.4). All sequences were aligned using Clustal W within the BioEdit program, version 7.1.8 (Hall, 1999) and examined for pairwise differences using version 5.10 (Librado and Rozas, 2009).

Morphological Analyses

As noted in the introduction, reassessment of the morphology of *A. corti*, *A. geminum* and *A. kenti* was in order. To connect molecular sequence data to morphological descriptions, efforts were made to obtain specimens from the type hosts and/or type localities, though this was not always possible. In particular, attempts to sample the type host of *A. corti* [tadpole madtoms, *Noturus gyrinus* (Mitchill, 1817)] were unsuccessful due to the extirpation of madtoms from Lake Mendota, WI. There have been no reports of tadpole madtoms in Lake Mendota since 1916 (Lyons 1989). Furthermore, efforts to sample madtoms within water bodies neighboring Lake Mendota were unsuccessful. I was able to resample the type localities and type hosts associated with *A. geminum* (Oneida Lake, NY; brown bullhead, *Ameiurus nebulosus* Lesueur, 1819) and *A. kenti*. In the case of *A. kenti*, type locality was not explicitly given in the original description, but rather listed as the lower Tallahatchie River (Simer 1929). The type locality based on the US National Parasite Collection (USNPC) records of the type and paratype specimens was Money, MS. Thus, I sampled at three locations on the Tallahatchie River within the Money, MS town limits: at Money Bayou, which connects

to the Tallahatchie (33° 39.493' N 90° 12.528' W), as well as within the Tallahatchie River itself both North (33° 43.120' N 90° 13.174' W) and South (33° 35.6874' N 90° 11.6990' W) of the bridge that crosses the river in the town.

In addition to the specimens collected by the authors for morphometric analyses, type and voucher specimens previously deposited in the US National Parasite Collection were obtained and measured. These specimens included *A. corti* (*n* = 20; type: USNPC 60211; vouchers: 97766, 95491, and 104447-104452 which were tied to sequence data by Tkach and Mills 2011), the type and paratype specimens of *A. kenti* (*n* = 2; USNPC 51371), the resurrected *A. kenti* specimens from Tkach and Mills (2011; *n* = 7; USNPC 104445-104446), and the type specimen of *A. geminum* (USNPC 008564). All measurements were taken from mounted specimens and are given in micrometers. These measurements are used to illustrate the key morphological differences that can help distinguish among the species.

Results

Molecular Sequence Data

2605 base pairs (bp) were originally aligned as part of the ribosomal DNA dataset, with 50 nucleotides subsequently removed due to associations with indel regions where alignment was ambiguous (Table 3). Pairwise *p*-distances and variable sites were calculated for the remaining 2555 bp fragments (18s and ITS1, Table 4; ITS2 and 28s, Table 5).

A. kenti, sampled from channel catfish (the type host) in Money, MS (the type locality) were found to be genetically distinct from the specimens deemed *A. kenti* by

Table 4. Percent pairwise *p*-distance (number of variable sites given in parenthesis) of the ITS1 region (508-base pair [bp] alignment; above the diagonal) and 18S gene (309-bp alignment; below diagonal) of nuclear ribosomal DNA among *Alloglossidium* species using ictalurid final hosts. Distances and number of variable sites were calculated after removing portions of the alignment that could not be aligned do to indels (see Chap 2, Table 3).

	A. corti	*A. kenti*	*A. geminum*	*A. progeneticum*	*A. floridense*	*A. fonti* (VT & WI)	*A. fonti* (LA & TX)	*A. n. sp.* A (TX & AR)	*A. n. sp.* A (NY)	*A. n. sp.* B
A. corti		2.8 (14)	5.1 (26)	1.6 (8)	1.4 (7)	2.4 (12)	2.4 (12)	2.2 (11)	2.2 (11)	2.8 (14)
A. kenti	0.6 (2)		5.5 (28)	2.6 (13)	2.6 (13)	3.7 (19)	3.7 (19)	3.0 (15)	3.0 (15)	2.0 (10)
A. geminum	0.6 (2)	0		4.9 (25)	4.9 (25)	5.9 (30)	5.9 (30)	4.3 (22)	4.3 (22)	5.1 (26)
A. progeneticum	0.6 (2)	0	0		1.2 (6)	2.6 (13)	2.6 (13)	2.0 (10)	2.0 (10)	2.6 (13)
A. floridense	0.6 (2)	0	0	0		2.4 (12)	2.4 (12)	2.0 (10)	2.0 (10)	2.6 (13)
A. fonti (VT & WI)	0.6 (2)	0	0	0	0		0.2 (1)	2.8 (14)	2.8 (14)	3.3 (17)
A. fonti (LA &	1.0 (3)	0.3 (1)	0.3 (1)	0.3 (1)	0.3 (1)	0.3 (1)		0.4 (2)	2.6 (13)	3.3 (17)
A. n. sp. A (TX)	0.6 (2)	0	0	0	0	0	0.3 (1)		2.6 (13)	2.6 (13)
A. n. sp. A (NY)	0.6 (2)	0	0	0	0	0	0.3 (1)	0		2.6 (13)
A. n. sp. B	0.6 (2)	0	0	0	0	0	0.3 (1)	0	0	

Table 5. Percent pairwise *p*-distance (number of variable sites given in parenthesis) of the ITS2 region (239-base pair [bp] alignment; above the diagonal) and about 1590 bases at the 5'end of the 28S gene (below diagonal) of nuclear ribosomal DNA among *Alloglossidium* species using ictalurid final hosts. Distances and number of variable sites were calculated after removing portions of the alignment that could not be aligned do to indels (see Chap 2, Table 3).

	A. corti	*A. kenti*	*A. geminum*	*A. progeneticum*	*A. floridense*	*A. fonti* (VT & WI)	*A. fonti* (LA & TX)	*A. n. sp.* A (TX)	*A. n. sp.* A (NY)	*A. n. sp.* B
A. corti		2.1 (5)	5.4 (13)	2.1 (5)	2.1 (5)	2.1 (5)	1.7 (5)	1.7 (4)	1.7 (4)	2.9 (7)
A. kenti	0.8 (13)		5.4 (13)	0.8 (2)	0.8 (2)	0.8 (2)	0.4 (1)	0.4 (1)	0.4 (1)	1.7 (4)
A. geminum	2.8 (44)	2.4 (38)		5.4 (13)	5.4 (13)	5.4 (13)	5.0 (12)	5.0 (12)	5.0 (12)	6.3 (15)
A. progeneticum	0.6 (9)	0.5 (8)	2.6 (42)		0.8 (2)	0.8 (2)	0.4 (1)	0.4 (1)	0.4 (1)	1.7 (4)
A. floridense	0.6 (10)	0.6 (9)	2.7 (43)	0.3 (5)		0.8 (2)	0.4 (1)	0.4 (1)	0.4 (1)	1.7 (4)
A. fonti (VT & WI)	0.9 (14)	0.8 (13)	3.0 (47)	0.6 (9)	0.6 (10)		0.4 (1)	0.4 (1)	0.4 (1)	1.7 (4)
A. fonti (LA & TX)	0.8 (12)	0.7 (11)	2.8 (45)	0.4 (7)	0.5 (8)	0.1 (2)		0	0	1.3 (3)
A. n. sp. A (TX)	0.5 (8)	0.5 (8)	2.6 (42)	0.3 (4)	0.3 (5)	0.4 (7)	0.3 (5)		0	1.3 (3)
A. n. sp. A (NY)	0.5 (8)	0.5 (8)	2.6 (42)	0.3 (4)	0.3 (5)	0.6 (9)	0.4 (7)	0.1 (2)		1.3 (3)
A. n. sp. B	0.9 (14)	0.4 (9)	2.6 (42)	0.6 (9)	0.6 (10)	0.9 (14)	(12)	0.6 (9)	0.6 (9)	

Tkach and Mills (2011); a comparison of the ribosomal sequences associated with these individuals found a combined ITS *p*-distance of 1.9%). Results of the mitochondrial sequence comparisons identified pairwise *p*-distances between the 2 "*kenti*" taxa of between 7.2-12.9% (Chapter II; Table S1). With these clear differences in sequences, congruence between independent genetic markers, and non-sister phylogenetic relationships (Fig. 7), I hereby delineate "*A. kenti*" to refer those specimens from the type locality (Money, MS) and "*Alloglossidium n. sp.* A" to refer to those specimens that share the rDNA sequence used by Tkach and Mills to resurrect the species.

Ribosomal DNA sequences were generally found to be conserved within species. However, intraspecific sequence variation was found among the *Alloglossidium n. sp.* A and *A. fonti* specimens. For *Alloglossidium n. sp.* A, 4 nucleotide (nt) substitutions (ITS1: A/T, G/A; 28s: G/T, G/A) were found between the New York individuals and all other specimens, regardless of geographic distribution (Tkach and Mills' 2011 sequences, Arkansas, and Texas) (Table 1; Fig. 7). Thus, *Alloglossidium n. sp.* A can be delineated into southern (Texas and Arkansas) and northern (New York) lineages (Table 4 and 5). For *A. fonti*, all 4 of the specimens from the current study have a T for the C/T substitution noted by Tkach and Mills (2011). Additionally, 2 substitutions were noted: specimen FO01 (Vermont) had an A/C substitution in the 18s and both specimens FO03 (Louisiana) and FO04 (Texas) had a T/G substitution in the 28s (Table 1; Fig. 7). Specimen F02 from Wisconsin was found to be identical to Tkach and Mills' previously reported sequences from North Dakota and Wisconsin. As such, the pairwise *p*-distances and number of variable sites are subdivided between geographical subsets

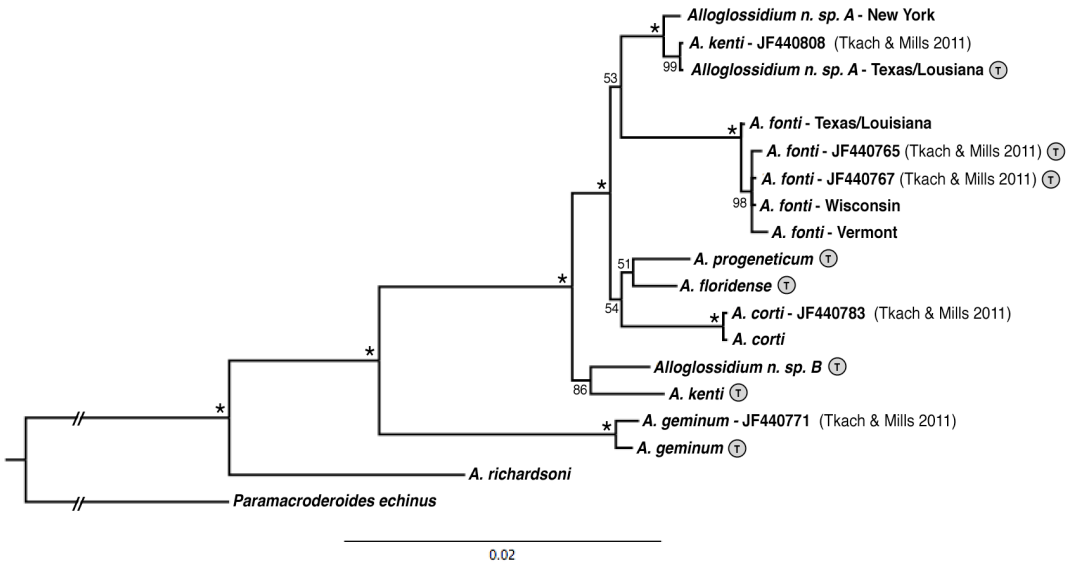

Figure 6. 50% majority rule consensus tree (18s-ITS1-5.8s-ITS2-28s rDNA) of *Alloglossidium* species from fishes. Numbers at nodes correspond to the Bayesian posterior probability percentages; * denote nodal support values of 100. The outgroup taxa *P. echinus* has been collapsed and branch length shortened. The outrgroup taxa *A. richardsoni* represents the clade of *Alloglossidium* that obligately uses leech definitive hosts. Grey circles denote sequences tied to the type locality and/or species description.

associated with each species. *Alloglossidium fonti* can also be separated into northern (Vermont and Wisconsin) or southern (Louisiana and Texas) groupings (Table 4 and 5). Interspecific variability based on *p*-distance was greatest in the ITS1 and 28s gene regions of the rDNA sequence (Table 4 and 5). *A. geminum* was the most highly diverged species with overall pairwise *p*-distances ranging between 3.0% and 3.5%. *Alloglossidium* n. sp. B was found to be most recently diverged from *A. kenti*, with an overall rDNA *p*-distance of 0.78% (20 bp differences).

Among mitochondrial haplotypes (as reported from Chapter II analyses), *p*-distances were found to be < 5%, with a few notable exceptions (Table S1). In the case of the *A. fonti* specimens, three individuals (FO01, Vermont; FO02, Wisconsin; and FO04, Texas) had *p*-distances between 2.2-3.3% when compared against each other, but pairwise comparisons with the fourth individual (FO03, Louisiana) yields *p*-distances of 5.5-5.9%, Additionally, among the *Alloglossidium n. sp.* A specimens, the 3 individuals from Texas (TK01, TK02) and Arkansas (TK03) differed by <1%, though the 2 individuals from New York (TK04 and TK05) showed differences of 6.8-7.7% (Table S1).

Morphological Data

An overview of 27 morphometric traits compiled from both newly surveyed specimens and the reassessment of specimens obtained from museum collections is provided in Table 6. The museum specimens were chosen due to their status as type and/or paratype individuals (*A. corti*, *A. kenti*, and *A. geminum*) or due to their use in the recent molecular studies (Tkach and Mills 2011). Morphological measurements and

characterization of the specimens lead to the distinction of 9 morphological traits that when taken as a comprehensive suite of characters, appear to be the most useful with regards to identifying species. These traits (i.e. body length and width, cirrus sac appearance, anterior and posterior extent of the vitellaria, ovary shape, ovary placement, and egg length and width) among the species that infect fish are outlined in Table 7. This table also includes comparisons to the species from fish for which complete morphological and molecular descriptions exist (*A. fonti*, Tkach and Mills 2011; *A. floridense*, Kasl et al. 2014) or recently reanalyzed (*A. progeneticum*, Kasl et al. 2015). Finally, line drawings of the 5 species measured in this study are included as a visual reference to highlight the subtle interspecific variation (Fig. 8: (A) *Alloglossidium n. sp.* B, (B) *A. corti*, and (C) *A. geminum*; Fig. 9: (A) *A. kenti* and (B) *Alloglossidium n. sp.* A).

Discussion

Molecular data were pivotal in recognizing cryptic diversity within *Alloglossidium* species from fishes; a combination of rDNA and mtDNA sequence data strongly support the designation of 8 species (and possibly a further 2 species). First, *A. fonti* (Tkach and Mills 2011), *A. floridense* (Kasl et al. 2014), and *A. progeneticum* (Kasl et al. 2015) are already well grounded, linking sequence data to complete morphological descriptions. As a result of the analyses reported herein, I found strong support that *A. kenti* from Mississippi is distinct from *Alloglossidium n. sp.* A (the specimens reported in Tkach and Mills 2011). Vilas et al. (2005) note that many morphologically recognized pairs of congener digeneans exhibit about 1% divergence in the ITS regions, which is the

Table 6. Morphometrics with mean values ± standard deviation (range) of adult *Alloglossidium* specimens from ictalurid fishes. Note that formal species descriptions for *Alloglossidium n. sp.* A and B will be addressed in a separate publication and as such the measurements included should not be considered as a formal diagnosis.

Measurements (µm)	*Alloglossidium corti*	*Alloglossidium kenti*	*Alloglossidium geminum*	*Alloglossidium n. sp.* A	*Alloglossidium n. sp.* B
Sample size (*n*)	20	17	19	24	11
Body length	1222.8 ± 401.7 (527-1875)	1918.9 ± 362.2 (1275-2500)	1860.5 ± 346.4 (1275-2350)	2654.6 ± 713.6 (1575-4100)	2568.2 ± 745.0 (1625-4125)
Body width at ventral sucker	247.9 ± 76.7 (101-425)	380.5 ± 117.7 (172-725)	329.4 ± 63.5 (233-450)	384.8 ± 94.7 (225-600)	253.3 ± 80.6 (61-325)
Forebody	334.2 ± 65.0 (191-444)	463.2 ± 100.1 (323-725)	517.8 ± 79.2 (364-627)	547.5 ± 144.0 (374-830)	601.7 ± 77.5 (485-707)
Oral sucker length	108.5 ± 19.0 (76-140)	166.7 ± 111.2 (101-586)	102.4 ± 19.0 (67-141)	127.3 ± 27.9 (71-219)	136.5 ± 33.9 (71-182)
Oral sucker width	122.0 ± 23.5 (71-160)	179.9 ± 39.7 (101-253)	131.3 ± 25.3 (91-182)	159.6 ± 38.0 (111-250)	169.1 ± 40.6 (91-232)
Prepharynx length	53.0 ± 33.0 (20-170)	41.3 ± 22.3 (10-81)	74.6 ± 21.6 (40-111)	65.3 ± 38.3 (20-162)	70.9 ± 26.9 (20-121)
Pharynx length	64.1 ± 23.1 (30-120)	85.0 ± 23.0 (51-141)	81.5 ± 16.1 (51-101)	83.4 ± 21.9 (51-141)	78.0 ± 15.7 (61-101)
Pharynx width	68.0 ± 26.3 (15-110)	103.1 ± 23.7 (81-172)	83.9 ± 21.9 (61-131)	91.0 ± 19.1 (61-125)	88.0 ± 14.2 (61-111)
Esophagus	21.3 ± 11.8 (10-61)	27.6 ± 15.0 (10-71)	53.3 ± 20.4 (20-91)	46.3 ± 25.2 (20-131)	31.4 ± 15.8 (10-61)
Caecal bifurcation to anterior end	263.3 ± 57 (130-340)	311.9 ± 53.3 (212-394)	340.3 ± 48.8 (263-444)	335.3 ± 86.2 (232-520)	312.4 ± 83.1 (156-424)
Ventral sucker length	78.4 ± 23.6 (37-120)	118.8 ± 27.2 (81-192)	100.5 ± 12.7 (81-121)	125.2 ± 30.1 (80-172)	120.0 ± 39.1 (81-192)
Ventral sucker width	77.9 ± 24.2 (37-121)	121.8 ± 27.3 (101-212)	101.8 ± 8.2 (81-121)	127.6 ± 37.0 (71-212)	106.3 ± 30.8 (51-152)
Cirrus sac length	169.1 ± 43.8 (94-280)	336.9 ± 77.9 (202-474)	245.1 ± 47.7 (152-354)	323.1 ± 77.4 (222-525)	238.1 ± 44.8 (162-283)
Cirrus sac width	39.0 ± 14.7 (20-70)	94.6 ± 26.1 (61-152)	61.3 ± 11.8 (41-91)	62.7 ± 28.0 (30-121)	50.7 ± 15.6 (30-71)
Genital pore to anterior end	320.2 ± 14.7 (141-424)	418.4 ± 76.5 (283-525)	495.9 ± 73.6 (364-606)	499.7 ± 125.3 (333-808)	532.8 ± 109.5 (354-697)
Vitelline field to anterior end	229.8 ± 45.3 (111-300)	350.5 ± 67.9 (222-495)	619 ± 99.7 (467-768)	433.3 ± 126.8 (273-750)	464.6 ± 128.1 (303-657)
Vitelline field to posterior end	498.6 ± 167.5 (217-800)	265.9 ± 729.6 (212-1142)	610.5 ± 154.4 (384-859)	1217.6 ± 445.0 (450-2175)	1224.2 ± 494.2 (707-1818)
Ovary length	116.0 ± 38.6 (66-200)	160.5 ± 62.7 (111-394)	110.0 ± 12.4 (81-131)	205.1 ± 47.3 (111-283)	192.8 ± 31.5 (151-253)
Ovary width	97.3 ± 25.2 (61-152)	146.8 ± 31.9 (101-222)	106.3 ± 19.3 (71-141)	178.0 ± 33.9 (121-242)	187.8 ± 26.2 (131-222)
Ovary to ventral sucker	37.3 ± 18.4 (10-71)	88.2 ± 36.8 (15-141)	140.4 ± 43.5 (61-212)	143.9 ± 60.1 (40-273)	113.6 ± 51.5 (10-172)
Anterior testis length	126.0 ± 42.6 (64-240)	203.7 ± 51.8 (131-293)	161.5 ± 28.3 (121-212)	274.5 ± 75.9 (141-414)	233.8 ± 57.9 (131-333)
Anterior testis width	112.7 ± 38.3 (60-180)	190.6 ± 45.9 (141-333)	134.4 ± 32.5 (71-202)	224.8 ± 53.8 (141-328)	203.4 ± 29.7 (172-253)
Posterior testis length	139.2 ± 40.4 (75-220)	254.2 ± 73.1 (162-404)	182.3 ± 33.2 (131-232)	347.7 ± 105.9 (192-578)	261.0 ± 55.4 (156-354)
Posterior testis width	121.4 ± 37.5 (52-182)	196.9 ± 50.2 (121-343)	150.4 ± 28.8 (101-202)	220.7 ± 55.1 (152-343)	200.8 ± 44.4 (111-253)
End of caeca to posterior end	89.0 ± 50.5 (20-222)	232.7 ± 120.8 (61-544)	126.0 ± 46.7 (71-222)	243.3 ± 146.5 (61-520)	432.5 ± 221.3 (222-848)
Egg length	26.4 ± 3.0 (20-33)	24.7 ± 3.4 (16.2-30)	29.8 ± 2.7 (25-35)	24.5 ± 2.4 (19-30)	27.0 ± 2.6 (20-30)
Egg width	14.2 ± 2.1 (10-18)	13.3 ± 2.3 (8.1-17.5)	17.2 ± 1.9 (15-20)	13.1 ± 1.8 (10-15)	14.7 ± 1.8 (10-20)

Table 7. The most useful morphological traits (μm) for identification of *Alloglossidium* species reported from ictalurid catfishes. With the exception of *A. corti* (which could not be obtained from its type locality), all traits reported herein are associated with molecular sequence data also tied to the type locality. Ranges and means (in parentheses) are provided. Note that, though included, *Alloglossidium n. sp.* A and B are not yet formally described.

Characters	*A. corti*	*A. kenti*	*A. geminum*	*A. progeneticum*		*A. floridense*	*A. fonti*	*Alloglossidium n. sp. A*	*Alloglossidium n. sp. B*
				obligate 3-host	facultative precocious				
Reference	this study	this study	this study	Kasl et al. 2015	Kasl et al. 2015	Kasl et al. 2014	Tkach and Mills 2011	this study	this study
Body length	527-1,875 (1276.5)	1275-2,500 (1918.9)	1,275-2350 (1,860.5)	825-2,500 (1724.1)	889-2,300 (1446.1)	687-1,263 (977)	1,005-1,820 (1,356.2)	1,575-4,100 (2654.6)	1,625-4,125 (2568.2)
Body width	101-425 (247.9)	172-725 (380.5)	233-450 (329.4)	175-465 (314.5)	202-950 (320.5)	126-288 (197)	167-311 (243.6)	225-600 (384.8)	161-325 (262.4)
Cirrus sac	Long, thin	Large, distended	Long, thin	Long, thin	Long, thin	Long, thin	Long, thin	Long, thin	Long, thin
Anterior extent of the vitellaria	Between pharynx and cecal bifurcation	Cecal bifurcation	Posterior margin of ventral sucker or further posteriorly	Just below the cecal bifurcation	Just below the cecal bifurcation	Just anterior to the genital pore to the anterior end of the ventral sucker	Anterior margin of ventral sucker or slightly lower	Cecal bifurcation to genital pore	Just anterior to the ventral sucker
Posterior extent of the vitellaria	Anterior margin of posterior testis	Anterior to posterior margin of anterior testis	Between the testes to the anterior margin of the posterior testis	Anterior margin of posterior testis	Anterior margin of posterior testis	Posterior margin of the anterior testis	Posterior margin of anterior testis or midway between testes	Anterior to posterior margin of anterior testis	Between the testes to the anterior margin of the posterior testis
Ovary shape	Ovoid	Lobed	Ovoid	Ovoid	Ovoid	Ovoid	Ovoid	Ovoid/Lobed	Ovoid
Ovary placement	Immediately postcirrus	Immediately postcirrus	Immediately postcirrus	Immediately postcirrus to <half the ovary diameter	Immediately postcirrus to <half the ovary diameter	Postcirrus by <half the ovary diameter	Postcirrus by ~half the ovary diameter	Immediately postcirrus	Immediately postcirrus to <half the ovary diameter
Egg length	20-33 (26.4)	16-30 (24.7)	25-35 (29.8)	14-24 (21.5)	18-25 (21.9)	25-30 (27)	22-27 (25.1)	19-30 (24.5)	20-30 (27.0)
Egg width	10-18 (14.2)	8-17.5 (13.3)	15-20 (17.2)	8-12 (10.5)	9-13 (11.2)	13-17 (14)	9-14 (11.8)	10-15 (13.1)	10-20 (14.7)

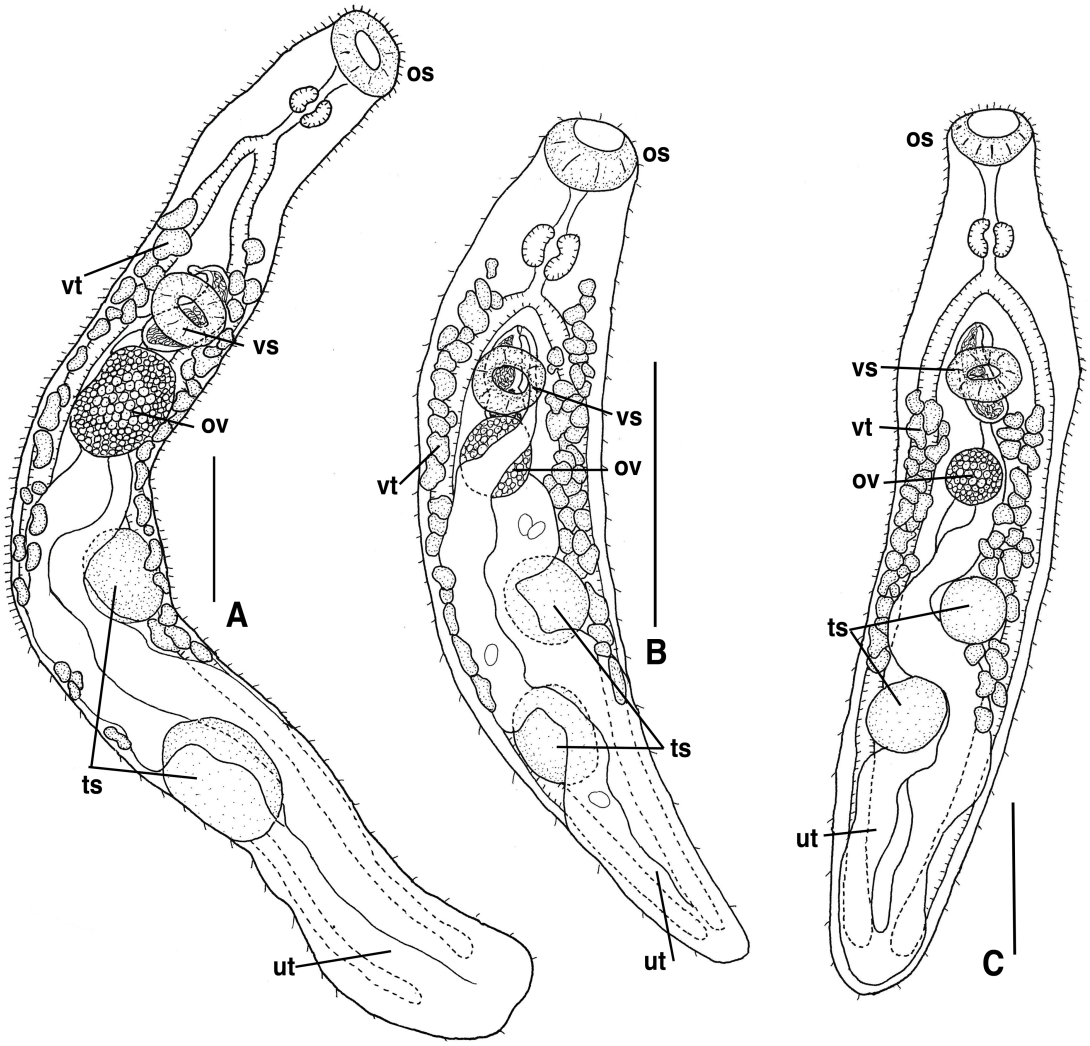

Figure 7. Ventral view of adult *Alloglossidium* spp. from fish. **A**) Ventral view of *Alloglossidium n. sp.* B, **B**) ventral view of *Alloglossidium corti*, and **C**) ventral view of *Alloglossidium geminum*. Locations of oral sucker (os), vitellaria (vt), ventral sucker (vs), ovary (ov), testes (ts), and uterus (ut) are indicated. Scale bars = 300 μm.

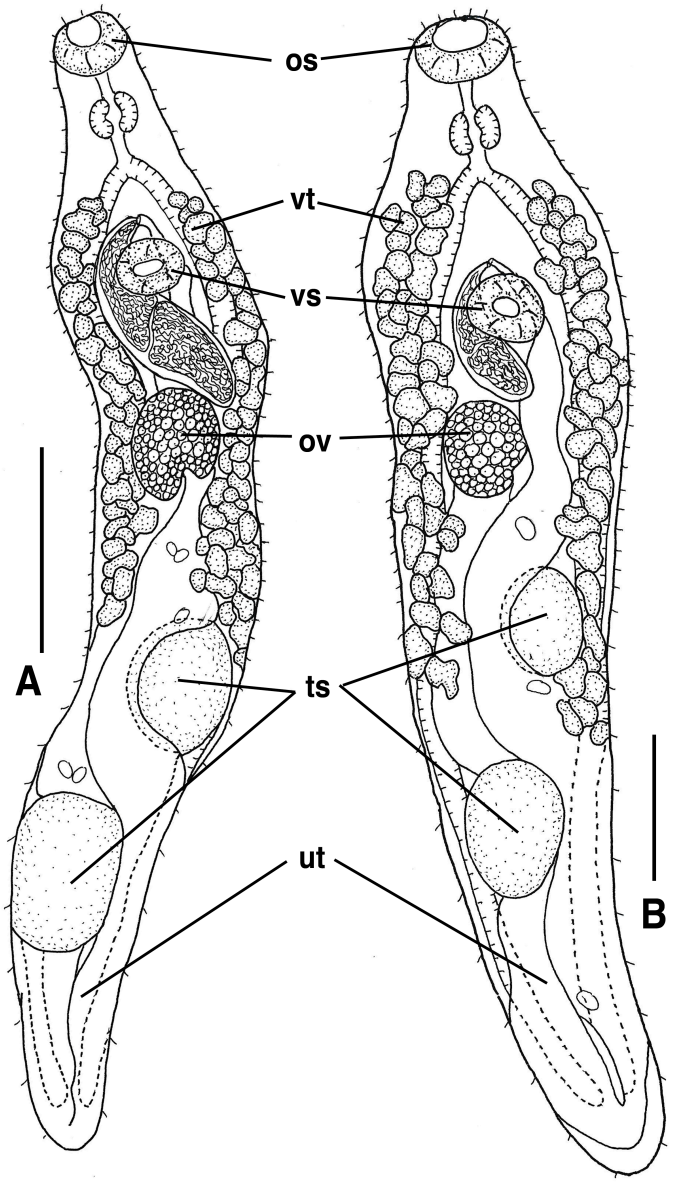

Figure 8. Ventral view of adult *Alloglossidium* spp. from fish. **A**) Ventral view of *Alloglossidium kenti* and **B**) ventral view of *Alloglossidium n. sp. A*. Locations of oral sucker (os), vitellaria (vt), ventral sucker (vs), ovary (ov), testes (ts), and uterus (ut) are indicated. Scale bars = 300 μm. Note that *Alloglossidium n. sp.* A was originally used to resurrect *A. kenti* by Tkach and Mills (2011).

approximate level of the variation between *A. kenti* and *Alloglossidium* n. sp. A (combined ITS *p*-difference = 1.9%; Table 4 and 5).

The situation here with *A. kenti* raises an important caution with trying to ascribe sequence data to a previously described species. *Alloglossidium kenti* was originally described by Simer (1921, but later synonymized with *A. corti* by Van Cleave and Mueller (1934). Tkach and Mills had collected samples from Reelfoot Lake (Lake Co., TN and Chotard Lake (Issaquena Co., MS) around the Tallahatchie system but not the type locality itself. They resurrected *A. kenti* due to genetic data that clearly delimited their samples from existing sequence data of nominal species. However, implicit in their resurrection was the assumption that their samples were representative of those in the type location. It was not until a comparison was made between the various locations that the cryptic variation between what turned out to be 2 species (*A. kenti* and *Alloglossidium n. sp. A*) became obvious. I acknowledge from my own experience (see below) that it is not always possible to collect from the type locality/type host for molecular studies and thus, the next best thing is to sample from locations in reasonable proximity to the type location (as done by Tkach and Mills). Nevertheless, it must be advised that a definitive link of a species description with sequence data is only possible if the samples are representative of the type location/host

In the case of *A. corti,* I returned to the type location (Lake Mendota, WI) to sample, but only *A. fonti* was recovered from brown and yellow bullheads. The extirpation of the type host, tadpole madtoms, from this location unfortunately means samples from its type locality can no longer be obtained. Therefore, it becomes even

more important to set a molecular-morphological species "benchmark." This is particularly true when dealing with the case of *A. corti*. Because cryptic species were unknown, *A. corti* was a "catch-all" taxon from 1935 to 2011. It is quite likely that the historical records are rife with incorrectly identified species and should be reexamined in light of the revised morphological characteristics to determine the validity of individual records. I was able to morphologically "ground" sequences associated with both *A. geminum*, which was resampled from its type host and locality, and the new species *Alloglossidium n. sp.* B.

Finally, the ND1 mtDNA sequence data indicate that there may be additional cryptic diversity within lineages of *A. fonti* and *Alloglossidium n. sp. A*. though further morphological analyses are still needed. In the case of *A. fonti*, there appears to be a north/south split, with rDNA sequences from Louisiana and Texas (LA & TX) distinct from those sampled from Wisconsin and Vermont (WI & VT). However, the mtDNA patterns are not completely congruent with the LA haplotype more similar to the WI and VT samples (Chapter II). This could represent a case of incomplete lineage sorting. The highly diverged specimens of *Alloglossidium n. sp.* A from New York provide a much more clear-cut case of cryptic diversity because the rDNA and mtDNA divergence patterns are congruent among the samples. However, additionally sampling is necessary to obtain enough specimens to thoroughly examine the morphology of those individuals.

With the above designations linking species identifications with sequence data, I examined the morphological traits of the 8 species known to infect fish. Overall, the morphology has very subtle variations, often with ranges that overlap between taxa.

The remeasurment of those traits currently used in species diagnoses provides a more complete record of the variation found in *A. corti*, *A. kenti*, and *A. geminum* then has previously been available.

Many of the traits identified as potentially useful in species identifications are found to overlap between different subsets of these taxa (Table 5). For instance, body size can be useful if a particular individual is quite large (*Alloglossidium n. sp.* A or B) or small (*A. corti* or *A. floridense*). While most *Alloglossidium* species in fish have an ovoid ovary, *A .kenti* has a lobed ovary, and *Alloglossidium n. sp.* A appears to have either an ovoid or lobed ovary shape. This clarification of ovary shape also serves to resolve a documented discrepancy between the original species description of *A. kenti*, which includes a line drawing of an ovoid ovary, and the genus diagnosis (based off of *A. kenti*), which states a lobed ovary as characteristic of the genus (Simer 1929). One possible explanation for this discrepancy may be the dark staining of the holotype specimen; in my own examination of the holotype, the ovoid shape of the ovary only became apparent under high levels of illumination. This ovary shape may be what led to the original designation of *Alloglossidium n. sp. A* as a *A. kenti* by Tkach and Mills (2011). However, *A. kenti* is on average a slightly smaller worm with a comparatively larger cirrus sac. Of all of the traits examined the most universally useful character appears to be the combined extent of the vitellaria. Though, I note that the determination of the start/end of the vitellaria is subjective and can be difficult to distinguish. Nonetheless, these results highlight that these traits must not be analyzed singly; overlap

between species suggests that no 1 trait in isolation can be considered diagnostic (Table 5).

Taken together, these findings highlight the need for formal redescriptions of *A. corti* (especially in regards to providing a resource of disambiguating 80 years of historical records), *A. kenti*, and *A. geminum* as well as descriptions of *Alloglossidium n. sp. A* and *Alloglossidium n. sp. B*. Providing complete morphological and molecular information for these species will establish a much needed reference to assist with the proper identification of *Alloglossidium* species from fish hosts. Additionally, creating such a reference would allow for the reanalysis of historical records to try and determine the host specificity and geographic ranges associated with the *Alloglossidium* species from fishes. These findings, along with recent taxonomic studies of the genus, illustrate how subtle morphological variation is an impediment to the conclusive identification of unique species, either by failing to recognize new species (Tkach and Mills 2011; Kasl et al. 2014) or by failing to recognize intraspecific variation (Kasl et al. 2015).

.

CHAPTER IV

EVOLUTIONARY CONSEQUENCE OF A CHANGE IN LIFE CYCLE COMPLEXITY: A LINK BETWEEN PRECOCIOUS DEVELOPMENT AND EVOLUTION TOWARD FEMALE-BIASED SEX ALLOCATION IN A HERMAPHRODITIC PARASITE*

Introduction

The complexity of parasite life cycles, which often involve diverse routes of infection and multiple host species to complete development, have long garnered interest among biologists. Theoretical studies have largely focused on the causes that led to the evolution of complex life cycles (Choisy et al. 2003; Parker et al. 2003; Gandon 2004). Empirical work has mainly focused on how parasite life history traits such as timing of reproduction, optimal growth patterns, or host manipulation may evolve as a response to changes in complex life cycle patterns (Gemmill et al. 1999; Kozlowski 1992; Hammerschmidt et al. 2009; Parker et al. 2009; Benesh et al. 2012; Benesh et al. 2013). Some studies have proposed that complex life cycles of hermaphroditic parasites may have evolved (or are maintained) as a need to encounter mates to promote outcrossing (Brown et al. 2001) or to avoid mating among identical clones in the case of trematodes (Rauch et al. 2005). However, a change in life cycle pattern itself may alter the parasite

---

* Reprinted with permission from "Evolutionary consequence of a change in life cycle complexity: A link between precocious development and evolution toward female-biased sex allocation in a hermaphroditic parasite" by EL Kasl, CT McAllister, HW Robison, MB Connior, WF Font, and CD Criscione, 2015. Evolution, 69, 3156-3170, Copyright 2015 by John Wiley & Sons Inc.

mating system and thus have downstream evolutionary consequences. In particular, among digenean trematodes some species may have altered mating systems due to the early onset of sexual maturation in an intermediate host (Lefebvre and Poulin 2005).

The most common digenean life cycle follows an obligate 3-host pattern. Miracidia that develop from eggs infect a molluscan first intermediate host. Asexual reproduction within the mollusc leads to the release of larval cercariae that subsequently infect a second intermediate host and encyst as metacercariae (a juvenile stage that is not yet sexually mature). Completion of the life cycle is attained upon ingestion of the second intermediate host by a definitive host (typically a vertebrate) wherein the parasite becomes sexually mature (Fig. 1A). However, there are notable variations in this life cycle pattern. In particular, there are digenean species that precociously reproduce (historically termed progenetic) in what is typically regarded as the second intermediate host (Lefebvre and Poulin 2005). For example, in the genus *Alloglossidium*, some species have a typical 3-host pattern with ictalurid catfishes as definitive hosts (where within the intestine, adult worms are free to outcross) whereas others have 2-host patterns with a crustacean or leech definitive host (Smythe and Font 2001). Of particular interest is *Alloglossidium progeneticum*, which has a facultative 2-host or 3-host life cycle. In this species the mating system is significantly altered by the precocious life cycle change because within its crayfish second host, worms become sexually mature within their cyst (Fig. 1B). Thus, this developmental change leads to forced self-fertilization. This drastic change in the mating system (increased selfing) may then have

downstream evolutionary consequences such as inbreeding depression or changes to reproductive traits themselves.

In this study, I focus on a potential evolutionary change in sex allocation as a consequence of a change in life cycle complexity that alters the parasite's mating system. In self-compatible hermaphroditic species, local sperm competition (competition between related sperm) is a concept that provides predictions for sex allocation in relation to the mating system (reviewed in Schärer 2009). In particular, with a high selfing rate or a small mating group size, sex allocation should favor a female-biased allocation, as an individual should only produce enough sperm to fertilize the available eggs. Schärer (2009) notes that the evolution of sex allocation may occur two ways: 1) It may occur as a fixed (or average within population) evolutionary response optimized over generations. I refer to this as a genetically-based response among populations. And 2) Sex allocation may occur as a phenotypically plastic response (individuals adjust sex allocation in response to current conditions). The few empirical studies on sex allocation in parasitic flatworms have focused exclusively on the latter (Trouve et al. 1999; Schärer et al. 2001; Schjørring 2009; Al-Jahdali 2012).

Here, I provide a test of an among population genetically-based response in sex allocation by comparing populations of *A. progeneticum* that vary in their life cycle pattern (obligate 3-host vs. facultative precocious 2- or 3-host life cycle) while controlling for potential environmental confounding factors. Specifically, I test the hypothesis that sex allocation evolved towards a more female-biased function in the derived facultative precocious populations. This study was made possible due to recent

survey efforts that discovered life history variation among populations. Thus, I first provide a brief summary of the newly discovered life history, geographic, and genetic variation in the trematode *A. progeneticum*. I then present statistical analyses of the reproductive morphology that show strong support for reduced allocation to male resources in populations with precocious development. These results emphasize how changes in life cycle patterns that alter mating systems can impact the evolution of reproductive traits in parasites.

Methods

Study Species

The digenetic trematode *A. progeneticum* is a simultaneous hermaphrodite that can facultatively use either 2- or 3-hosts to complete development (Fig. 1B). Originally described from gravid specimens encysted within the antennal glands of the crayfish *Procambarus spiculifer* from Calls Creek, Georgia (Sullivan and Heard 1969), a subsequent study of *A. progeneticum* in Calls Creek found adults free in the intestines of brown bullheads, *Ameiurus nebulosus* (Font and Corkum 1975). I note, however, surveys only found snail bullheads, *Ameiurus brunneus*, in Calls Creek (Table S2). The presence of gravid specimens (thus sexually mature) within both what is typically considered the second intermediate host (crayfish) and definitive third host (ictalurids) is to date the only such occurrence within the genus *Alloglossidium* (Smythe and Font 2001; Kasl et al. 2014). Prior to this study *A. progeneticum* had only been reported from Calls Creek (Sullivan and Heard 1969; Font and Corkum 1975). Recent survey efforts,

reported herein, identified geographic variation in the ability to precociously develop, with most populations exhibiting an obligate 3-host life cycle pattern (Fig. 10, Table S2).

Collections

From June 2011 through May 2014 ictalurid catfishes and crayfish hosts were sampled across the southern U.S. (Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Alabama, Georgia; Fig. 10, Table S2). All parasites taken from the intestines of catfishes were heat-fixed (passing the wet-mount over the flame of an alcohol burner) underneath a cover slip without pressure to ventrodorsally orient the worms in preparation of morphological analyses (Font 1994). Cysts taken from crayfish hosts were examined for eggs under a compound microscope and classified as either encysted metacercariae (i.e., not gravid) or encysted adults (i.e., gravid). Some gravid cysts were mechanically excysted and heat-fixed as above to allow for morphological analyses. All parasites were stored in 70% ethanol to allow for potential future molecular analyses.

Extraction, DNA Amplification, and Sequencing

DNA was extracted by placing individual worms in 200 µl of 5% chelex containing 0.2 mg/ml of proteinase K, incubated for 2 hr at 56°C and boiled at 100°C for 8 min. Two markers were used for *A. progeneticum* identification and analysis of genetic diversity across its expanded range: an 839 base pair ribosomal DNA (rDNA) fragment spanning the 3' end of the 18s and the first internal transcribed spacer region (ITS1), and a 679 base pair region of the mitochondrial NADH-dehydrogenase subunit 1 gene (ND1). Polymerase chain reaction (PCR) amplifications were performed using 25 µl reactions. The nuclear ribosomal sequence was obtained using a reaction consisting

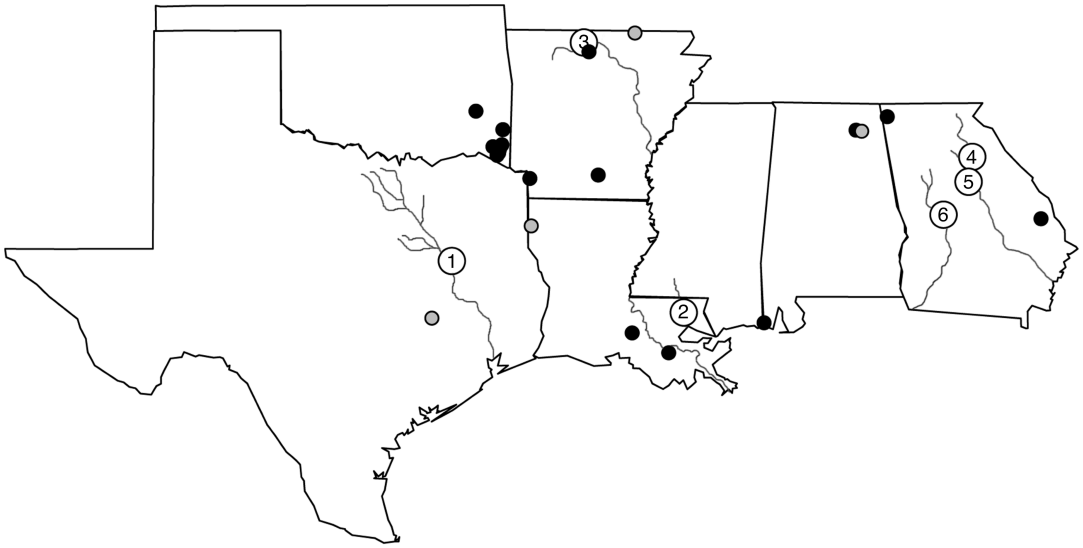

Figure 9. Sites positive for *Alloglossidium progeneticum* in the Southern U.S. confirmed with sequence data. Black dots represent sites where non-gravid metacercariae were recovered from crayfish hosts, thus indicating non-precocious development. Grey dots represent sites where only adult worms were collected from fish, thus precocious life history cannot be ascertained. Sites used for comparative analyses are indicated by circles with numbers. Sites 1 (Gus Engeling), 2 (Chappapeela Creek), and 3 (Crooked Creek) are obligate 3-host populations. Sites 4 (Calls Creek, type locality), 5 (Big Indian Creek), and 6 (Richland Creek) are the only identified facultative precocious populations in my survey.

of 3 µl of template extract, 16.25 µl water, 2.5 µl 10x buffer, 1.5 µl $MgCl_2$ [25 mM], 0.5 µl dNTP [10 mM/each], 0.5 µl of each primer [20 µM], and 0.25 µl of Taq polymerase (Omega Bio-Tek Inc., Norcross, GA) [5 units/µl], and a thermocycling profile of 95°C for 3 min, once; 94°C for 45 sec, 60°C for 30 sec, 72°C for 60 sec, 35 times; 72°C for 7 min, once. For the rDNA region I used the forward primer s18 (5'-TAACAGGTCTGTGATGCC-3'; located towards the 3' end of the 18s rDNA) and reverse primer 5.8s1 (5'-GCTGCGCTCTTCATCGACA-3'; located within the 5.8s gene) (Jousson et al. 2000). The s18 primer was used for sequencing. PCR for the ND1 region was the same with the exceptions that 15.25 µl water and 2.5 µl $MgCl_2$ [25 mM] were used. The thermocycler profile for the ND1 PCR is described in Criscione and Blouin (2004). The forward primer MB352 (5'-CGTAAGGGKCCTAAYAAG-3") (Criscione and Blouin 2004) and reverse primer CC28 (5'-CWTCTCAARGTTAACAGCCT-3'; anchored in the asparagine tRNA) were used for the PCR of the ND1 region. In addition to MB352, an *A. progeneticum*-specific internal reverse primer CC57 (5'-CCCATAATCTATGTGTGCTAAC-3') was used for sequencing. PCR products were purified using the Ultra Clean PCR clean-up Kit (MO BIO Laboratories, Inc., Solana Beach, California) and sent to the DNA Analysis Facility on Science Hill at Yale University (New Haven, Connecticut) for sequencing. Contiguous sequences from individuals were assembled using Sequencher™ (GeneCodes Corp., ver. 4.1.4), and submitted to GenBank under accession numbers KT455707-KT455827. Both nuclear and mitochondrial sequences were aligned using Clustal W within the BioEdit program, version 7.1.8 (Hall, 1999).

Molecular Identification of Specimens as A. progeneticum

For the survey work, at least one worm per host species was sequenced at each site positive for specimens of *Alloglossidium* (Table S2). Due to the subtle morphological differences among species of *Alloglossidium* known to infect catfishes (Kasl et al. 2014), molecular sequences from new sites were compared to those taken from the type locality (Calls Creek, GA) for positive identification of *A. progeneticum*. I also compared the rDNA region to sequences of other species in the genus *Alloglossidium* including Genbank sequences (Tkach and Mills 2011: JF440783.1, *A. corti*; JF440765.1, *A. fonti*; JF440771.1, *A. geminum*; JF440808.1, *A. kenti*; Kasl et al. 2014: KC812276.1, *A. floridense*) and sequences from additional known or cryptic species obtained as part of my ongoing survey work (to be published elsewhere as part of a phylogenetic study of the genus). For the ND1 sequences, TCS version 1.21 (Clement et al. 2000) was used to construct a haplotype network. This network is not intended to be a final representation of the phylogeographic patterns as sampling was haphazard and opportunistic. Rather the purpose is to support the status of the specimens as *A. progeneticum* and to provide initial data on intraspecific variation especially in the 6 focal populations used for the morphological analyses (discussed below). ND1 haplotype (H_d) and nucleotide diversity (π) within these 6 focal populations was calculated in DnaSP version 5.10 (Librado and Rozas, 2009).

Morphological Analyses to Test for Sex Allocation

Based on populations with suitable sample sizes, the molecular confirmation of populations as being *A. progeneticum,* and the presence of encysted metacercariae or

encysted gravid adults, I chose 6 representative populations for downstream analyses. Calls Creek, Big Indian Creek, and Richland Creek (Fig. 10; sites 4, 5, and 6, respectively) are classified as facultative precocious (2 or 3-host life cycle) populations due to the presence of encysted gravid worms in crayfish and gravid worms in bullheads. More detail on my survey findings is given in the Results, but basically all encysted worms in crayfish from these populations are gravid. Additionally, Richland Creek is a different river drainage (Flint River system) than that of Calls Creek and Big Indian Creek (Oconee River system). Gus Engeling, Chappapeela Creek, and Crooked Creek (Fig. 10; sites 1, 2, and 3 respectively) are classified as obligate 3-host populations as encysted metacercariae in crayfish from these locations were not gravid. The obligate 3-host populations represent 3 non-connected river systems.

All specimens used for morphological analyses were stained using Semichon's acetic carmine, dehydrated in a graded ethanol series, cleared using xylene, and mounted on slides using either Damar gum or Canada balsam. Because the morphology of *A. progeneticum* has not been described from adult specimens originating from fish hosts, especially from obligate 3-host populations, I provide a list of morphological measurements for traits commonly used in trematode taxonomy in Table S3.

To assess if the life cycle change leading to precocious development ultimately leads to an evolutionary change in sex allocation, I compared reproductive traits within and between the 6 representative populations discussed above. Ideally, I would use volume measurements of the reproductive organs. However, the staining and mounting methods, which were necessary for specimen identification, did not permit measurement

of volume. As surrogates of volume, I measured the length and width of the ovary and both testes. Length was the longest axis going from anterior to posterior and width was the longest axis perpendicular to length (see Fig. 11). An exploratory principle components analysis (PCA with Varimax rotation in SYSTAT v12; SYSTAT Software, Inc., Chicago, IL) of all measured worms revealed that the 6 measures of the reproductive organs fell out as two factors. One factor was composed of the ovary length and width and the other factor was composed of the lengths and widths of both testes (Table S4). Based on this PCA I created summated scores to represent female function [(ovary length + ovary width)] or male function [(anterior testis length + anterior testis width + posterior testis length + posterior testis width)]. Summated scales are an appropriate way to summarize correlated variables of the same trait (Hair et al. 1998). Sex allocation was estimated as a ratio: male function/(male function + female function) (Janicke and Schärer 2009; Janicke et al. 2013). As noted by Janicke et al. (2013), this estimator is a relative measure of sex allocation where higher values indicate more male-biased sex allocation. This estimator also carries the assumption that testis and ovary size are reasonable proxies for investment into male and female sex function, respectively. Additional discussion of appropriate means to measure sex allocation can be found in Schärer (2009).

In the following statistical analyses I employed generalized linear mixed models (GLMM). For all analyses, the sex allocation ratio of male function/(male function + female function) was used as the dependent variable. Following the recommendations of Schärer (2009), if significance was observed, I repeated the same analyses on male or

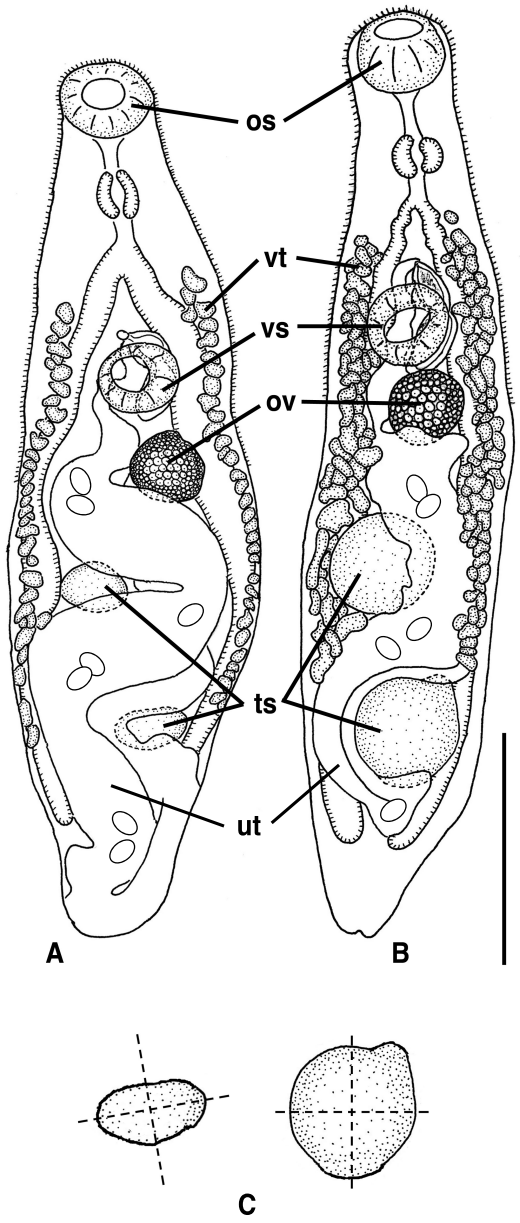

Figure 10. Ventral view of adult *Alloglossidium progeneticum* specimens from fish. **A**) Worm from a precocious population (Calls Creek, GA) and **B**) worm from a non-precocious population (Gus Engeling, TX). Locations of oral sucker (os), vitellaria (vt), ventral sucker (vs), ovary (ov), testes (ts), and uterus (ut) are indicated. For simplicity, the uterus shows a few representative eggs but is typically completely filled with eggs. Scale bar = 400 μm. Note testes are much smaller in the specimen from the facultative precocious population (A) than the obligate 3-host population (**B**). **C**) Panel showing how length (vertical dotted line) and width (horizontal dotted line) were measured.

female function separately to determine if changes in one or both of these traits were driving sex allocation. All statistical analyses were performed using the package lme4 (to fit restricted maximum likelihood models (REML); Bates and Maechler, 2009) and lmerTest (to obtain *P*-values using the Satterthwaite approximation; Kuznetsova et al. 2015) in the program R (R Core Team, 2015). Residuals from all of the tests were visually examined to see if the data met the assumptions of normality and homogeneity of variance.

Because selection on sex allocation may favor a phenotypically plastic response to current conditions (Schärer 2009), an analysis that tests for a genetically-based evolutionary response among populations can be confounded if specimens are collected from different conditions (e.g., high vs low intensity of infection). Thus, I first conducted statistical tests within the obligate 3-host populations and within the facultative precocious populations to look for evidence of plastic responses. Within the obligate 3-host populations Gus Engeling and Crooked Creek, I had sufficient samples from fish to test for the influence of intensity of infection (number of worms per infected host) on sex allocation. The premise here is that with a lower intensity in fish, local sperm competition will be higher (i.e., there is a smaller social group size; Schärer 2009). Thus, my estimate of sex allocation should be greater under higher intensities. Parasites were separated into low intensity (i.e., originating from a fish with < 20 worms) or high intensity (i.e., originating from a fish with >20 worms) categories to evaluate the plastic adjustment of sex allocation. This threshold was chosen to correspond to the highest observed intensity from the precocious populations (see

below). I measured at least 10 worms for each category from each of the two populations (total $N = 45$). Worms came from at least 2 different host individuals in each category per population. The model included the fixed effect of intensity, body size (total length of the worm) as a covariate, and the interaction between the two. The random effects included population, the interaction between population and intensity, and host individual nested within the population by intensity interaction. I did not include body width along with total length of the worm as summated score of body size because I found body width to be unreliable due to the fact that heavily gravid worms are expanded in width.

Within all 3 of the facultative precocious populations I only had adult worms from fish that had infection intensities < 20. However, adults are also present as encysted worms within crayfish. Thus, I could test for the influence of an overall outcrossing opportunity on sex allocation by comparing worms from fish, which are freely able to outcross, and crayfish, wherein encysted worms are forced to self-mate. The idea here is that with forced self-fertilization, sex allocation would shift towards a female-biased function because there are no additional mates besides oneself (i.e., there is high local sperm competition). From each of the 3 populations, I measured at least 10 worms for each of two categories: forced self-fertilization in crayfish vs opportunity to outcross in fish (total $N = 63$). For simplicity, I refer to this classification as the mating opportunity fixed effect. The model included the fixed effect of mating opportunity, body size as a covariate, and the interaction between the two. The random effects

included population, the interaction between population and mating opportunity, and host individual nested within the population by mating opportunity interaction.

Based on the results of the within population analyses above, I limited the among population analysis (i.e., the test for a genetically-based evolutionary response in sex allocation as a result of the change in the life cycle) to a comparison of adult worms obtained from low intensity fish infections (originating from a fish with < 20 worms). The goal here was to test the hypothesis that the increased self-fertilization (i.e., an increase in local sperm competition) in the facultative precocious populations would lead to a female-biased shift in sex allocation. The analysis included all 6 of the representative populations and included 10 worms from each population. Worms came from at least 2 different fish in each population (total $N = 62$). The model included the fixed effect of life history (obligate 3-host vs facultative precocious), body size as a covariate, and the interaction between the two. The random effects included population nested within life history and host individual nested within population within life history.

Results

Survey Results: Life History, Morphological, and Molecular Variation

In total, 28 locations (including the type locality of Calls Creek) were found to have *A. progeneticum* (verified with sequence data, discussed below). Encysted adults in crayfish were found only in 3 sites in Georgia (Calls Creek, Big Indian Creek and Richland Creek; sites 4, 5, 6 in Fig. 10, respectively). In these sites, all encysted worms are gravid except for the very small ones, which likely indicate a recent infection where the worm has not yet had time to mature. Moreover, gravid worms were also recovered

from the intestines of ictalurids in sites 4, 5, and 6 (Fig. 10, Table S2). Thus, these 3 populations are considered facultative precocious populations. From 4 populations, I only recovered worms from fish hosts, thus I cannot definitively classify the life history of these populations (Fig. 10, Table S2). Of the other 21 sites where infected crayfish were found (N = 151 infected crayfish), I recovered 1352 encysted metacercariae that were of comparable size to those found in sites 4, 5, and 6 discussed above. None of these metacercariae were gravid (Table S2). Thus, I classify these latter 21 sites as obligate 3-host populations as a fish host would be necessary for the worms to mature.

Comparisons of rDNA sequences obtained through survey efforts (N = 119) to those taken from the type locality (Calls Creek; site 4, Fig. 10) led to the identification of 27 new locations of *A. progeneticum*, expanding the known range across 6 states (Fig. 10; Table S2). Moreover, 9 new species of crayfish and 5 new ictalurid species were identified as intermediate and final hosts, respectively (Table S2). Within the ITS1 region of the rDNA sequence, a 14-nucleotide insert (5'-TTATCCTAAAAGGT-3') is diagnostic compared to this region of the 5 other described species of *Alloglossidium* that infect fish (Kasl et al. 2014). Among the *A. progeneticum* specimens, there was a single variable site that had no association with geography, host species, or life history. As this site occurred at the end of a mononucleotide repeat (8 or 9 repeats of A), its validity remains suspect. At ND1 (N = 119), 37 unique haplotypes were identified (Fig. 12; Table S2). No premature stop codons were found after translation of the sequences (using amino acid translation code 9 on GenBank). Maximum p-distance between any

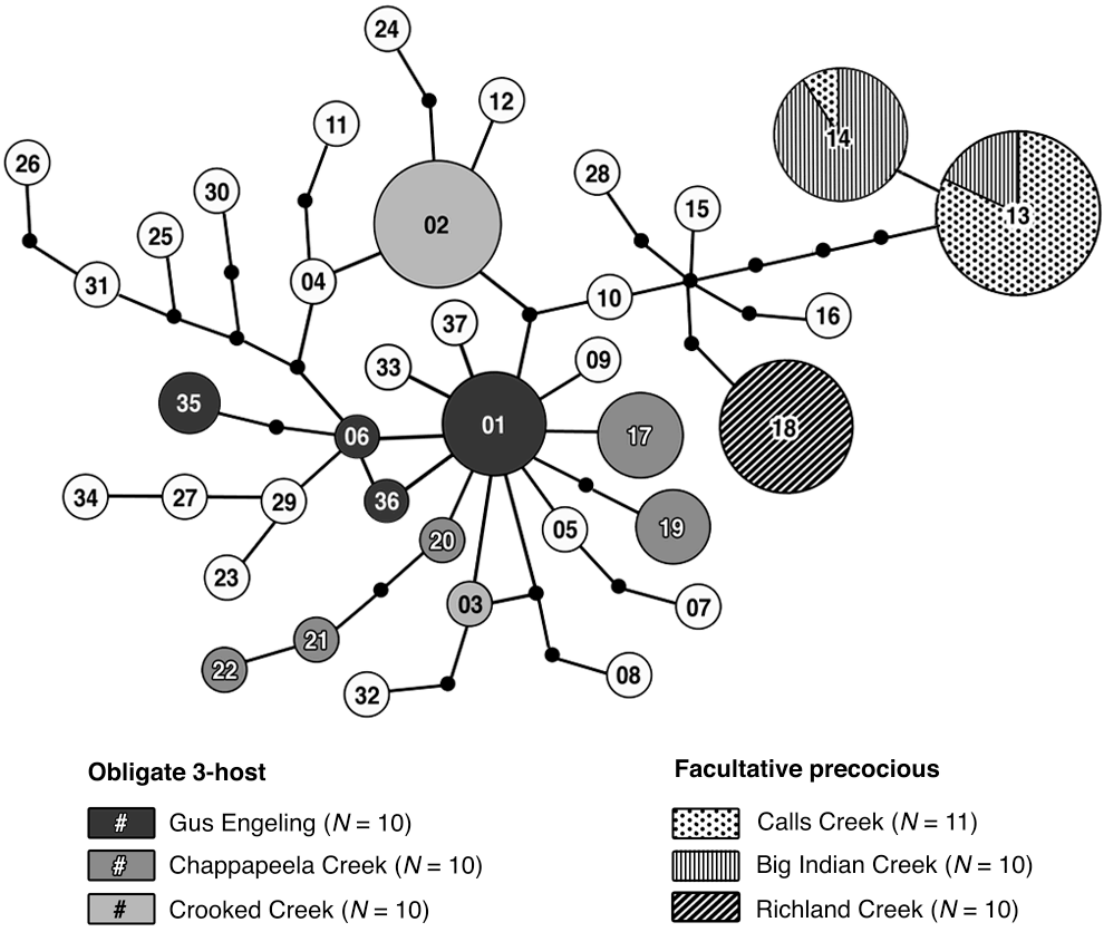

Figure 11. Statistical parsimony network of ND1 mtDNA haplotypes. Numbers represent haplotype ID (Table S2). Each connection is a single mutational step with small black circles representing inferred haplotypes. The network is intended to show the relationships and diversity of haplotypes collected from the 6 populations used for the morphological analyses. Haplotypes from these 6 focal populations are shown as shaded (obligate 3-host) or patterned (facultative precocious) circles and are scaled according to the number of individuals with that haplotype (e.g. 1 individual from Gus Engeling had haplotype 6). The network is not intended to be a final representation of phylogeographic patterns due to opportunistic surveys resulting in uneven sampling. White haplotypes are those not represented in the 6 focal populations and are not scaled. Haplotypes from the obligate 3-host populations may be shared with other non-focal sampling locations, but this is not shown in the figure (see Table S2). However, haplotypes from the facultative precocious populations are not shared with any of the non-focal sites nor the obligate 3-host populations.

two ND1 haplotypes was 1.6%, which is well within the range typically found for intraspecific mtDNA variation among parasitic flatworms (Vilas et al. 2005).

Of the 37 haplotypes identified, only 3 corresponded to the facultative precocious populations. Of these, 2 haplotypes (haplotype 13 and 14) were shared between sites within the same river drainage (Calls Creek and Big Indian Creek, Oconee River system) while the third (haplotype 18) was unique to the Richland Creek site, located within the Flint River drainage (Fig. 12; Table S2). Hence, haplotype (H_d) and nucleotide diversity (π) is low within these facultative precocious populations ($H_d = 0.2$, $\pi = 0.0003$; $H_d = 0.33$, $\pi = 0.0005$; $H_d = 0$, $\pi = 0$; Calls Creek, Big Indian Creek, and Richland Creek, respectively). In general, the obligate 3-host populations had greater haplotype and nucleotide diversity ($H_d = 0.8$, $\pi = 0.0042$; $H_d = 0.64$, $\pi = 0.0021$; $H_d = 0.2$, $\pi = 0.0009$; Chappepeela Creek, Gus Engeling, and Crooked Creek, respectively).

Infection Intensity and Plasticity in Sex Allocation

For populations with the obligate 3-host life cycle, when treating sex allocation as the response variable the interaction between intensity and body size (total length of the worm) was not significant ($F_{1,41} = 0.75$, $P = 0.39$). Thus, the interaction was pooled. After pooling, the covariate, total worm length, had a significant positive relationship with sex allocation (i.e., the larger the worm, the more allocation to the male function; $F_{1,42} = 13.58$, $P < 0.001$; Fig. 13A). Intensity of infection was also significant where the high intensity category corresponded to higher male-biased sex allocation ($F_{1,42} = 9.59$, $P = 0.003$; Fig. 13B). In the analyses of the individual summated scales of male and female function, interactions between intensity and body size were not significant (male

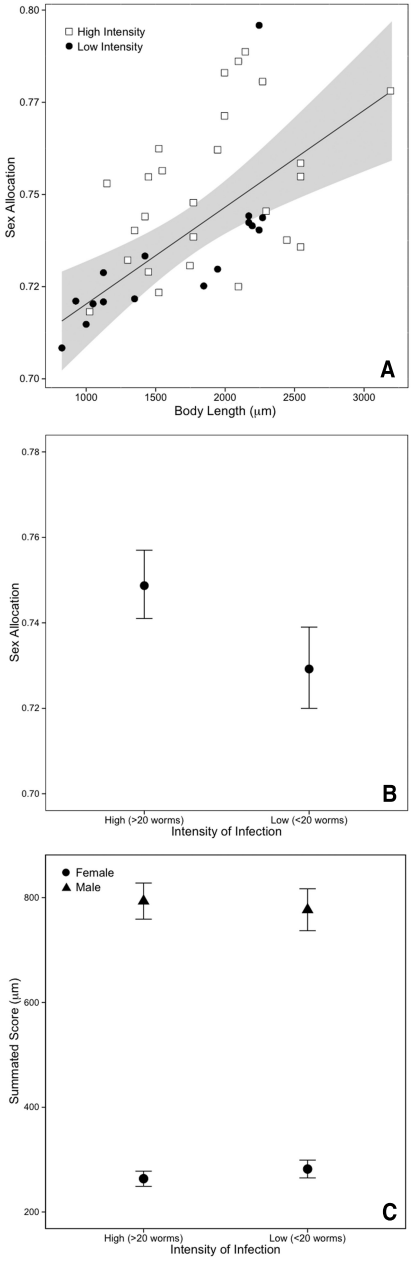

Figure 12. Linear mixed model results for the fixed effect of infection intensity for worms from 2 of the obligate 3-host populations of *Alloglossidium progeneticum* (Gus Engeling and Crooked Creek). **A**) Relationship of body size (μm) to sex allocation. Sex allocation is expressed as the ratio of the male summated scores to the combined male and female summated scores. Shaded regions represent 95% confidence intervals based on the model standard errors. The raw data points (i.e., not model adjusted) are superimposed over the model prediction. **B**) Least square means plot showing the relationship between intensity of infection and sex allocation. **C**) Least square means plot showing the relationship between intensity of infection and individual female (ovary) and male (combined testes) summated scores (μm). Error bars in **B** and **C** represent 95% confidence intervals.

function: $F_{1,41} = 2.94$, $P = 0.09$, female function: $F_{1,41} = 1.85$, $P = 0.18$). After pooling interactions, body size had a significant positive association to both male ($F_{1,42} = 176.72$, $P < 0.001$) and female ($F_{1,42} = 63.04$, $P < 0.001$) functions. However, infection intensity was not significant for either male ($F_{1,42} = 0.38$, $P = 0.54$) or female ($F_{1,42} = 2.57$, $P = 0.12$) functions. Despite the lack of significance with infection intensity to male or female function alone, there was a general trend of slightly higher summated scales for female function in cases of low intensity and slightly higher summated scales for male function in cases of high intensity (Fig. 13C). Thus, it appears that the combined changes in male and female function are driving the overall sex allocation in the direction of greater male-biased function under higher intensities. Random effects did not account for a significant portion of the variance in any of the above 3 models (Table S5).

Mating Opportunity and Plasticity in Sex Allocation

For the test of mating opportunity (as a function of host type) on sex allocation within populations with a facultative precocious life cycle pattern, the interaction between mating opportunity and body length was significant ($F_{1,55} = 11.25$, $P = 0.001$, Fig. 14A). I conducted a post-hoc analysis (using R package lsmeans) to see where along the gradient of body size I started to see a statistical difference in sex allocation between worms from fish and crayfish. This occurs at a body size of 1696 μm (df = 2.21, *t*-ratio = -3.93, $P = 0.05$), so worms from fish with body sizes greater than 1696 μm have higher sex allocation than worms from crayfish. In addition, I conducted post-hoc analyses to determine if the slopes for worms from fish and crayfish were different

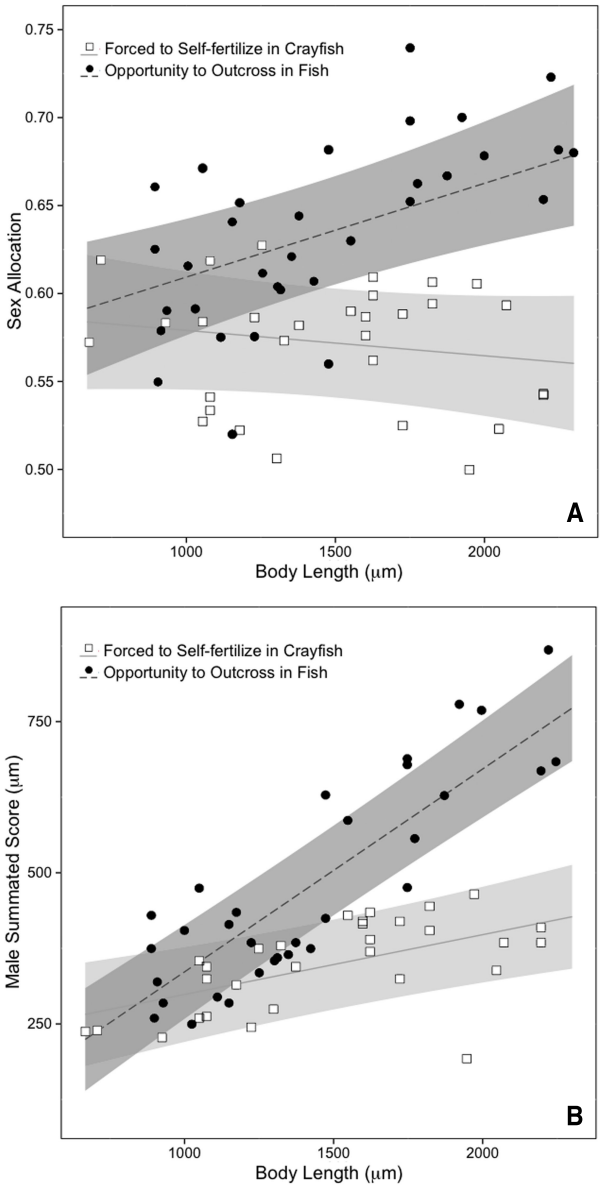

Figure 13. Linear mixed model plot showing the interaction between total worm length (μm) and mating opportunity (opportunity to outcross in fish versus forced self-fertilization in crayfish) on sex allocation (**A**) or on male summated scores (**B**) in sites with facultative precocious development (Calls Creek, Big Indian Creek, Richland Creek). Note in **A** sex allocation does not scale with body size when there are no potential mates other than oneself (i.e., encysted in crayfish host). Sex allocation is expressed as the ratio of the male summated score to the combined male and female summated scores. Shaded regions represent 95% confidence intervals based on the model standard errors. The raw data points (i.e., not model adjusted) are superimposed over the model predictions.

than 0. A significant positive relationship between body size and sex allocation was observed for worms collected from fish ($\beta = 5.37x10^{-5}$, $F_{1,30} = 19.69$, $P < 0.001$) whereas worms encysted in crayfish had a non-significant negative relationship between body size and sex allocation ($\beta = -1.21x10^{-5}$, $F_{1,28} = 0.61$, $P = 0.44$; Fig. 5A). When conducting the analyses on the female summated scales, the interaction was not significant ($F_{1,54} = 1.54$, $P = 0.22$). After pooling, body size had a significant positive association with female function ($F_{1,56} = 71.27$, $P < 0.001$), but mating opportunity was not significant with female function ($F_{1,19} = 0.99$, $P = 0.33$). In contrast, male function showed a significant interaction between mating opportunity and body size ($F_{1,56} = 39.19$, $P < 0.001$; Fig. 14B). As this result was similar to that found for sex allocation, it appears that changes in the male function are largely driving the significance found for sex allocation. In general the random effects did not account for a significant portion of the variance in the above 3 models (Table S6).

Life History and Genetically-based Response in Sex Allocation

Because the above results suggest a plastic response of sex allocation to the current conditions worms experience (e.g., infection intensity or mating opportunity), the tests for a genetically-based evolutionary response among populations with different life histories were restricted to adult worms obtained from fish with low intensities of infection (originating from a fish with < 20 worms). A significant interaction was found between life history and body length of the worm ($F_{1,56} = 5.17$; $P = 0.02$, Fig. 15A). Post-hoc analysis of sex allocation across the range of body sizes showed that there was a difference in sex allocation between life histories (greater in the obligate 3-host until a

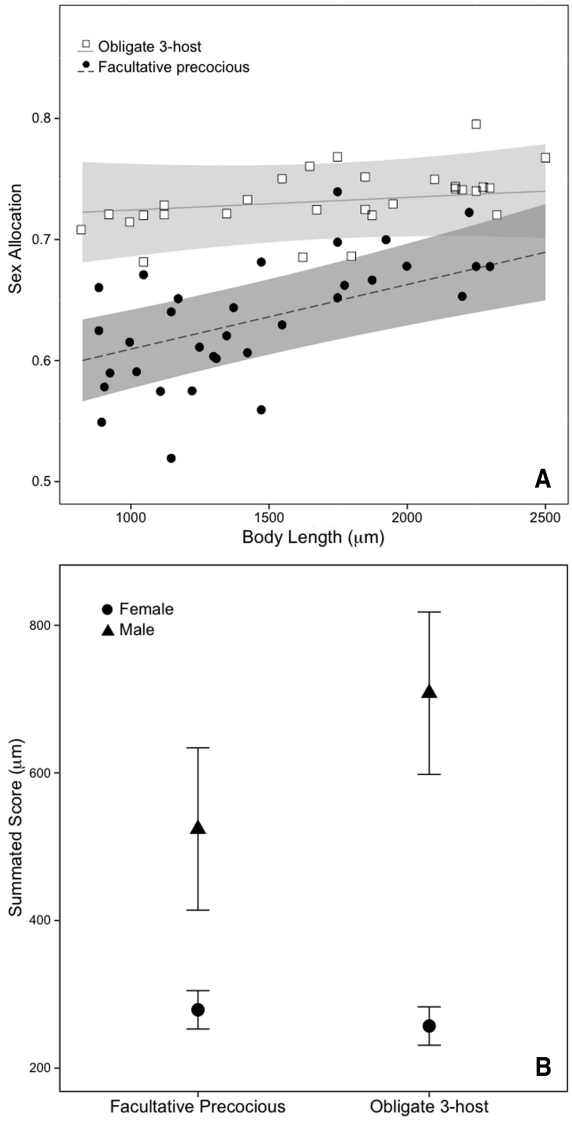

Figure 14. Results of the life history (Facultative precocious vs. obligate 3-host life cycle pattern) linear mixed models. **A**) The interaction between life history and body size on sex allocation. Sex allocation is expressed as the ratio of the male summated score to the combined male and female summated scores. Shaded regions represent 95% confidence intervals based on the model standard errors. The raw data points (i.e., not model adjusted) are superimposed over the model prediction. Note, up to a body size of 2224 μm the obligate 3-host worms have significantly greater male biased sex allocation relative to the facultative precocious individuals (discussed in text). **B**) Least square means plot showing the relationship between life history and female (ovary) and male (combined testes) summated scores (μm). Error bars represent 95% confidence intervals.

size of 2224 μm, at which point there was no longer a difference (df = 5.85, *t*-ratio =2.46, $P > 0.05$). At the maximum-recorded size of 2500 μm non-significance was marginal (df = 8.37, *t*-ratio = 1.77, $P = 0.11$). Thus, over most of the range of body size, the obligate 3-host life history has a more male-biased sex allocation compared to the facultative precocious life history. With the latter in mind, the main fixed effects still could be interpreted. Body size had a significant positive association to sex allocation ($F_{1,56} = 11.43$, $P = 0.001$). Life history was significant where the facultative precocious life history showed more of a female-biased sex allocation relative to the obligate 3-host life history ($F_{1,19} = 17.5$, $P < 0.001$; Fig. 15A). Subsequent analyses of the female summated scores showed a non-significant interaction between life history and body size ($F_{1,23} = 0.46$, $P = 0.51$). After pooling, body size had a significant positive relationship with female function ($F_{1,26} = 70.4$, $P < 0.001$). Life history was not significant ($F_{1,4} = 2.74$, $P = 0.18$) though the trend was for greater female function in the facultative precocious life history (Fig. 15B). Male function also had a non-significant interaction ($F_{1,55} = 0.59$, $P = 0.45$). After pooling, both a positive relationship with body size ($F_{1,58} = 150.4$, $P < 0.001$) and life history ($F_{1,4} = 10.87$, $P = 0.03$) were significant. Male function was greater in the obligate 3-host life history (Fig. 15A). As with the mating opportunity analysis above, most of the change in sex allocation in the life history analysis appears to be driven by changes in male function. In the models testing sex allocation and male summated scores, random effects accounted for a significant portion of the variance (Table S7).

Discussion

Summary of Main Results

1. Survey data verified with sequence data expanded the geographic distribution (across 6 states in the southern U.S.), host species distribution (5 final ictalurid hosts and 9 intermediate crayfish hosts), and life history variation (obligate 3-host and facultative precocious) *of A. progeneticum*.

2. ND1 sequence data support the intraspecific status of specimens across locations and life history. However, haplotypes associated with facultative precocious populations were not shared with those found in obligate 3-host populations. Additionally, haplotypes were not shared between the facultative precocious populations in different river drainages.

3. Within obligate 3-host populations, worms taken from fish with a high intensity of infection had higher male-biased sex allocation than worms recovered from fish with low intensities. Worm body size also had a significant positive relationship with sex allocation.

4. Within facultative precocious populations, there was a significant interaction between worm body size and mating opportunity. In particular, there was a positive relationship between worm body size and sex allocation in worms from fish hosts (i.e., opportunity to outcross), but this relationship was absent in encysted adult worms from crayfish hosts (i.e., forced self-fertilization).

5. The test for a genetically-based response of sex allocation showed that the populations with the facultative precocious life cycle evolved towards a female-biased

sex allocation compared to the populations with an obligate 3-host life cycle. This pattern was primarily driven by a decrease in male function in the facultative precocious populations.

Elucidating the Biology of A. progeneticum

Prior to this study, the only known location of *A. progeneticum* was Calls Creek and the only known life history was that of a facultative precocious life cycle (Sullivan and Heard 1969; Font and Corkum 1975). My survey and sequence data provide several key insights into the ecology and evolutionary history of *A. progeneticum*. First, *A. progeneticum* is much more widespread than originally thought (Fig. 10). Second, there is variation among populations in the developmental patterns of its life history and third, several ictalurid and crayfish species can serve as hosts. Because of the subtle morphological differences among the species of *Alloglossidium* that infect ictalurids, it is likely that *A. progeneticum* specimens were lumped under the umbrella classification of other *Alloglossidium* species (especially *Alloglossidium corti*) in studies prior to the use of molecular data (Tkach and Mills 2011, Kasl et al. 2014). This lumping was also likely compounded by the fact that the variation in life cycle pattern went unrecognized. For example, reports of non-gravid, encysted metacercariae in crayfish were assumed to be *A. corti* (McAllister et al. 2011) based on historical life cycle work (McCoy 1928). Indeed, sequence data have since verified that the encysted metacercariae from crayfish that were identified as *A. corti* in McAllister et al. (2011) are in fact, *A. progeneticum*. Adult specimens from slender madtoms (*Noturus exilis*) reported in McAllister et al. (2015) are also *A. progeneticum*. The measurements provided in Table S3, along with

the extent of the vitellaria (extending from just below the cecal bifurcation to the anterior margin of the posterior testis; Fig. 11) can be used as a means to help with species identification in the absence of sequence availability (Kasl et al. 2014). Nonetheless, sequence data were, and will continue to be, imperative in correctly identifying new populations of *A. progeneticum*.

The identification of *A. progeneticum* from the other *Alloglossidium* species that infect fish is clearly discernable from the 14-nucleotide insert in the ITS1. The intraspecific status of my samples of *A. progeneticum* is supported by the lack of variation across this rDNA region and the low divergence at the mtDNA ND1 region. I will be addressing the phylogeny of the genus in a future manuscript. However, it is important to note here that preliminary findings based on these sequence data indicate the obligate 3-host life cycle is the ancestral trait and the facultative precocious life cycle is the derived life history in *A. progeneticum* (E. Kasl, W. Font, and C. Criscione unpublished).

The ND1 data are also suggestive of two other interesting features of *A. progeneticum* evolutionary history that will warrant further investigation. First, haplotype diversity and nucleotide diversity tend to be much lower in the facultative precocious populations (sites 4, 5, and 6; Figs. 10 and 12) compared to the populations with an obligate 3-host life cycle (sites 1, 2, and 3; Figs.10 and 12). Because all encysted worms in the facultative precocious populations are forced to self-mate, inbreeding is assumed to be higher in these populations. While inbreeding reduces the effective size (and hence measures of genetic diversity) at an autosomal locus, the

mating system should not impact the genetic diversity at a haploid locus like the ND1 (Graustein et al. 2002, Charlesworth 2003). Nevertheless, the biology of self-maters (e.g., a single individual may colonize a population) may promote bottleneck/founder events that will result in low genetic diversity (Charlesworth 2003). It may also be that a selfing-gene (i.e., a mutation that enabled maturation while in the cyst) has swept through these populations resulting in a bottlenecking of a progenitor mtDNA haplotype. Additional data are needed to disentangle the causes of reduced genetic diversity in the facultative precocious populations. Second, no haplotypes are shared between the Flint River drainage (haplotype 18 in Richland Creek) and Oconee River drainage in Georgia (haplotypes 13 and 14 in Calls Creek and Big Indian Creek) (Fig. 12). Moreover, the ND1 network (Fig. 12) suggests two possible origins of the evolution of facultative precocious development. For example, haplotype 18, which was only in Richland Creek, has less divergence to haplotypes 15 and 16 than to 13, which was in Calls and Big Indian Creeks (Fig. 12). Haplotypes 15 and 16 were from non-gravid metacercariae (i.e., obligate 3-host life history) collected from a nearby location in Georgia (Jackson Branch, Table S2). Additional loci will be needed to confirm if facultative precocious development has independently evolved among populations of *A. progeneticum*. If indeed, these data reflect independent origins, then the evolution towards female-biased sex allocation in the facultative precocious populations (as discussed below) would also reflect independent evolutionary events in these two drainages.

Phenotypic Plasticity of Sex Allocation

The effect of intensity of infection on sex allocation, in which worms from the high intensity category (i.e., larger social group size) had higher male-biased sex allocation, is in line with previous experimental findings on the plastic responses of sex allocation in the trematode *Echinostoma caproni* (Trouve et al. 1999) and the free-living flatworm *Macrostomum lignano* (Janicke et al. 2013). For instance, Trouve et al. (1999) found that an increase in the number of unique clones within a host increased allocation to male function and decreased allocation to female function. The above studies focused on the lab manipulation of mating group size. From field-based collections of the trematode *Gyliauchen volubilis*, Al-Jahdali (2012) reported greater male-biased sex allocation from higher intensity infections. Here, I also find support for mating group size (assumed here to correlate to social group size as measured by infection intensity) as a driver of a plastic response in sex allocation in a natural environment. I recognize that a caveat of my study is that I do not know how many unique clones (resulting from the asexual phase in the snail, Fig. 1) are present within an individual host even at high intensities. Nonetheless, population genetic studies on trematodes with similar life cycle patterns (e.g., obligate aquatic 3-host life cycles) have shown that very few repeated clones are present within the final fish host (Criscione and Blouin 2006; Criscione et al. 2011; Gorton et al. 2012). Thus, I reason it is a safe assumption that the number of genetically unique individuals, and hence social group size, increases with the intensity of infection.

With regards to the overall opportunity to outcross (free in a fish gut vs encysted in a crayfish), the interaction between body size and mating opportunity is an intriguing result. I observed a positive relationship between body size and sex allocation from worms in fish (i.e., there was greater allocation to the male function in larger individuals; Fig. 14A). This is concordant with the positive relationship of body size and sex allocation observed in the infection intensity analysis where all worms were from fish (see results section on the relationship between intensity of infection and sex allocation, Fig. 13A). In contrast, when worms were forced to self-mate, there was no relationship between body size and sex allocation (Fig. 14A). Thus, in the facultative precocious populations, there appears to be a plastic response for the size-dependent relationship to sex allocation that depends on the mating opportunity (forced self-fertilization in crayfish vs. opportunity to outcross in fish).

Size dependent sex allocation has been reported in simultaneous hermaphroditic organisms (see Schärer 2009). However, the relationship may be positive or negative. Theoretical work shows that a negative relationship would result when small individuals with few resources should invest in male function due to a steep increase in male fitness gain. As individuals get larger, sex allocation would become female biased because fitness from male function has diminishing returns (Schärer 2009). A positive relationship may result if there is a competitive advantage to gain more mates with a larger body size. With more mates, larger individuals would have more male-biased sex allocation (Schärer et al. 2001; Schärer 2009). In my case, I observed a positive relationship. If competitive ability is the explanation, then one would expect that if

competition is removed, then the positive relationship would no longer hold. Indeed, I see no body size relationship with sex allocation for the encysted adult worms recovered from crayfish hosts (i.e., there is no mate competition with forced self-mating) (Fig. 14A). Regardless of the explanation for the observed positive relationship, the interaction highlights another potential evolutionary consequence for a change in the mating system. That is, sex allocation itself may not be the only trait that evolves, but a size dependent relationship (or lack thereof) with sex allocation may also evolve.

Life History and Sex Allocation

We found strong support that the change in life history from obligate 3-host to facultative precocious development, which led to a change in mating system with potentially higher inbreeding, led to sex allocation that was more female-biased (Fig. 15A). This change largely resulted from a reduction in male function (Fig. 15B). The reduction of the testes size in the facultative precocious specimen is even clearly discernable in a line drawing comparison (Fig. 11). Even though the ovary size tended to be larger in the facultative precocious life history, there was no significant difference (Fig. 15B). I note, however, that other reproductive traits such as the vitellaria, which provide nourishment to the embryos, may also be important in terms of female allocation. The three dimensional overlapping and fragmentation of the vitellaria precluded an accurate measurement with my current staining and mounting methods.

To control for the potential plastic responses in sex allocation noted in the prior section, I limited my analysis on the genetically-based evolutionary response of sex allocation to adult worms sampled from fish hosts with low intensities of infection.

Because I controlled for these potential environmental influences, my results should reflect a genetically-based response in the evolution of a more female-biased sex allocation in the facultative precocious life history.

A potential caveat has to deal with the actual mating system itself. While I can be certain that there is forced-self mating in the crayfish hosts of the facultative precocious populations, I do not yet know if worms from fish in the obligate 3-host populations are outcrossing. However, population genetic studies on trematodes with similar life cycle patterns (e.g., aquatic obligate 3-host life cycles) have been shown to repeatedly have Hardy-Weinberg Equilibrium, and thus, are necessarily outcrossing (Criscione and Blouin 2006; Criscione et al. 2011; Gorton et al. 2012). In addition, cross-experiments have repeatedly shown that when trematodes are in the presence of other conspecifics they outcross (Nollen 1983; Agatsuma and Habe 1985; Trouve et al. 1999). Thus, I consider it is a safe assumption that worms from fish in the obligate 3-host populations at least have a higher level of outcrossing than compared to worms in the facultative precocious populations. Nonetheless, a direct test of the levels of inbreeding is warranted as it represents another potential consequence of a change in a parasite's life cycle pattern.

To my knowledge, my study represents the first to examine for a genetically-based among-population evolutionary response in sex allocation in parasitic flatworms. The change in sex allocation is clearly associated with a change in the parasite's mating system. In comparison to some other studies that have examined among-population variation in sex allocation in animals, I find support that changes in mating systems or

modes of reproduction can impact the evolution of sex allocation among populations. For example, Johnston et al. (1998) found an association between the mating system of a self-compatible hermaphroditic snail and sex allocation. Individuals from populations with higher selfing-rates (estimated from a single locus) devoted a lower proportion of reproductive tissue to sperm production. In the free-living flatworm *Schmidtea polychroa*, populations with a greater frequency of parthenogenesis as a mode of reproduction had reduced sperm production (Weinzierl et al. 1998).

Overall, I believe that the significance of the results reported herein lies in the fact that a change in the life cycle pattern can alter the mating system and thus have consequences for the evolution of parasite reproductive traits. Specifically, forced selfing vs opportunities for outcrossing can alter allocation to male or female function. There are several other species of digenean trematodes that have independently evolved life cycles that deviate from the 3-host pattern as a result of precocious development within what is typically regarded as an intermediate host (Lefebvre and Poulin 2005). Some of these species are also forced to self-fertilize as a result of maturation within a cyst. For instance, cryptic trematode species within *Stegodexamene anguillae* show among lineage variation in the frequency of precocious maturation while encysted in an intermediate host (Herrmann et al. 2014). It would be interesting to see if sex allocation towards female-biased function correlates to the frequency of forced self-fertilization in these closely related parasites. In a broader context, it will be worthwhile to compare across the digenean phylogeny to test if sex allocation towards female-biased function

evolves repeatedly in these different lineages where precocious life cycle changes have occurred.

CHAPTER V

CONCLUSIONS

The research presented in my dissertation establishes the evolutionary relationships of, highlights the need for taxonomic revision within, and elucidates an evolutionary consequence of life cycle change in the genus *Alloglossidium.* My body of work redefines this system as a unique model with which to address not only the evolutionary trajectory of complex life cycles, but also the downstream evolutionary consequences associated with those changes.

Chapter II provided new insight into the interrelationships of the genus *Alloglossidium*, setting up a molecular framework with which to independently analyze character trait evolution. Unlike previously proposed hypotheses relating to the evolution of life cycle complexity within the genus, our findings present strong support for multiple independent origins of precocious, 2-host life cycle patterns. Species maturing in leeches evolved from a single divergence event. Species utilizing a 2-host pattern maturing in crustaceans have evolved at least 3 times, represented by well supported clades; 1) *A. anomophagis*, which retains an encysted developmental stage, 2) *A. caridicolum* and *A. dolandi*, which have lost the encysted developmental stage; and 3) the "*progeneticum* clade" (Clade II), which itself possibly contains multiple unique origins of precociousness. Though only the evolution of life cycle complexity was addressed within this work, the updated phylogeny provides a useful framework with

which to address traits that may also be associated with life cycle changes (e.g., loss of musculature, degree of spination, and overall body shape).

Chapter III highlighted the subtle morphological (i.e., cryptic) variation present within the *Alloglossidium* species from ictalurid hosts. The use of molecular data to delimit species in turn helped establish the extent of intra- and inter-specific morphological variation. Ribosomal sequence data identified candidates for 2 new species, one of which was previously incorrectly resurrected as *A. kenti*. Furthermore, I addressed the need to redescribe *A. corti*, the true *A. kenti*, and *A. geminum*, to definitively tie together sequence data with the level of morphological variation that was either absent from or misrepresented (due to past identifications under the umbrella of *A. corti*) in their original species descriptions. Finally, these findings highlight the key morphological characteristics, to be used cumulatively, that are associated with *Alloglossidium* in fishes: body size, egg size, extent of the vitellaria, shape and placement of the ovary, and appearance of the cirrus sac.

Chapter IV discussed how altered mating systems (due to precocious sexual maturation in what is typically regarded as an intermediate host) may impact opportunities for outcrossing. In turn, reproductive traits evolved to optimize sex allocation. Specifically, I tested the hypothesis that sex allocation evolved towards a more female-biased function in populations of the hermaphroditic digenean trematode *A. progeneticum* that can precociously reproduce in their second hosts. In these precocious populations, parasites are forced to self-fertilize as they remain encysted in their second hosts. In contrast, parasites in obligate 3-host populations have more opportunities to

outcross in their third host. I found strong support that in populations with precocious development, allocation to male resources was greatly reduced. I also identified a potential phenotypically plastic response in a body size-sex allocation relationship that may be driven by the competition for mates. These results emphasize how changes in life cycle patterns that alter mating systems can indirectly influence the evolution of reproductive traits in parasites.

REFERENCES

Agatsuma, T. and S. Habe. 1985. *Paragonimus ohirai*: Genetic control of tetrazolium oxidase isozymes. Exp. Parasitol. 60: 309-313.

Al-Jahdali, M. O. 2012. Infrapopulations of *Gyliauchen volubilis* Nagaty, 1956 (Trematoda: Gyliauchenidae) in the rabbitfish *Siganus rivulatus* (Teleostei: Siganidae) from the Saudi coast of the Red Sea. Parasite 19: 227-238.

Bates, D. and M. Maechler. 2009. lme4: Linear mixed-effects models using S4 classes. available at: http://CRAN.R-project.org/package=lme4. Accessed 12 June 2015.

Beckerdite, F. W., and K. C. Corkum. 1974. *Alloglossidium macrobdellensis* sp. n. (Trematoda: Macroderoididae) from the leech, *Macrobdella ditetra* Moore, 1953. J. Parasitol. 60: 434-436

Benesh, D. P., J. C. Chubb, and G. A. Parker. 2013. Complex life cycles: why refrain from growth before reproduction in the adult niche? Am. Nat. 181: 39-51.

Benesh, D. P., F. Weinreich, and M. Kalbe. 2012. The relationship between larval size and fitness in the tapeworm *Schistocephalus solidus*: bigger is better? Oikos 121: 1391-1399.

Benard, M. F. and S. J. McCauley. 2008. Integrating across life-history stages: consequences of natal habitat effects on dispersal. Am. Nat. 171: 553-567.

Bergsten, J. 2005. A review of long-branch attraction. Cladistics 21: 163-193.

Bray, R. A., D. Gibson, and A. Jones [Eds.]. 2008. Keys to the Trematoda, Vol 3. CAB International and Natural History Museum, London.

Brown, S. P., F. Renaud, J.-F. Guégan, and F. Thomas. 2001. Evolution of trophic transmission in parasites: the need to reach a mating place? J. Evol. Biol. 14: 815-820.

Brooks, D. R. 2003. Lessons from a quiet classic. J. Parasitol. 89: 878–885.

Carney, J. P. and D. R. Brooks. 1991. Phylogenetic analysis of *Alloglossidium* Simer, 1929 (Digenea: Plagiorchiiformes: Macroderoididae) with discussion of the origin of truncated life cycle patterns in the genus. J. Parasitol. 77: 890-900.

Charlesworth, D. 2003. Effects of inbreeding on the genetic diversity of populations. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 358: 1051-1070.

Chauhan, V. and A. K. Pandey. 2014. Structure and evolution of the pod in *Indigofera* (Fabaceae) reveals a trend towards small thin indehiscent pods. Bot. J. Linn. Soc. 176: 260-276.

Choisy, M., S. P. Brown, K. D. Lafferty, and F. Thomas. 2003. Evolution of tropic transmission in parasites: why add intermediate hosts? Am. Nat. 162: 172-181.

Clement M., D. Posada, and K. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657-1660.

Corkum, K. C. and F. W. Beckerdite. 1975. Observations on the life history of *Alloglossidium macrobdellensis* (Trematoda: Macroderoididae) from *Macrobdella ditetra* (Hirudinea: Hirudinidae). Am. Midl. Nat. 93: 484-491.

Corkum, K. C. and Turner, H. M. (1977). *Alloglossoides caridicola* gen et sp. n. (Trematoda: Macroderoididae) from Louisiana crayfish. Proc. Helminth. Soc. Wash. 44: 176-178.

Crawford, W. W. 1937. A further contribution to the life history of *Alloglossidium corti* (Lamont), with especial reference to dragonfly naiads as second intermediate hosts. J. Parasitol. 23: 389-399.

Cribb, T. H., R. A. Bray, P. D. Olson, and D. T. J. Littlewood. 2003. Life cycle evolution in the Digenea: a new perspective from phylogeny. Adv. Parasitol. 54: 197-254.

Criscione, C. D., and W. F. Font. 2001. Artifactual and natural variation of *Oochoristica javaensis*: Statistical evaluation of in situ fixation. Comp. Parasitol. 68: 156-163.

Fish, T. D., and F. J. Vande Vusse. 1976. *Hirudicolotrema richardsoni* gen. et sp. n. (Trematoda: Macroderoididae) from Minnesota hirudinid leeches. J. Parasitol. 62: 899-900.

Criscione, C. D., and M. S. Blouin. 2004. Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. Evolution 58: 198-202.

Criscione, C. D., and M. S. Blouin. 2006. Minimal selfing, few clones, and no among-host genetic structure in a hermaphroditic parasite with asexual larval propagation. Evolution 60: 553-562.

Criscione, C. D., R. Vilas, E. Paniagua, and M. S. Blouin. 2011. More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. Mol. Ecol. 20: 2510-2524.

Font, W. F. and K. C. Corkum. 1975. *Alloglossidium renale* n. sp. (Digenea: Macroderoididae) from a fresh-water shrimp and *A. progeneticum* n. comb. T. Am. Microsc. Soc. 94: 421-424.

Font, W. F. 1980. The effect of progenesis on the evolution of *Alloglossidium* (Trematoda, Plagiorchiida, Macroderoididae). Acta Parasitol. Pol. 27: 173-183.

Font, W. F. 1994. *Alloglossidium greeri* n. sp. (Digenea: Macroderoididae) from the Cajun dwarf crayfish *Cambarellus schufeldti* in Louisiana, USA. T. Am. Microsc. Soc. 113: 86-89.

Feutry, P., M. Castelin, J. R. Ovenden, A. Dettai, T. Robinet, C. Cruaud, and P. Keith. 2013. Evolution of diadromy in fish: Insights from a tropical genus (*Kuhlia* species). Am. Nat. 181: 52-63.

Gandon, S. 2004. Evolution of multihost parasites. Evolution 58: 455-469.

Gemmill, A. W., A. Skorping, and A. F. Read. 1999. Optimal timing of first reproduction in parasitic nematodes. J. Evol. Biol. 12: 1148-1156.

Gorton, M. J., E. L. Kasl, J. T. Detwiler, and C. D. Criscione. 2012. Testing local scale panmixia provides insights into the cryptic ecology, evolution, and epidemiology of metazoan animal parasites. Parasitology 139:981-997.

Graustein, A., J. M. Gaspar, J. R. Walters, and M. F. Palopoli. 2002. Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. Genetics 161: 99-107.

Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black, eds. 1998. Multivariate Data Analysis. 5th ed. Prentice Hall, Upper Saddle River, New Jersey.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 4:95-98.

Hammerschmidt, K., K. Koch, M. Milinski, J. C. Chubb, and G. A. Parker. 2009. When to go: Optimization of host switching in parasites with complex life cycles. Evolution 63: 1976-1986.

Herrmann, K. K., R. Poulin, D. B. Keeney, and I. Blasco-Costa. 2014. Genetic structure in a progenetic trematode: signs of cryptic species with contrasting reproductive strategies. Int. J. Parasitol. 44: 811-818.

Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

Janicke, T., and L. Schärer. 2009. Determinants of mating and sperm-transfer success in a simultaneous hermaphrodite. J. Evol. Bio. 22: 405-415.

Janicke, T., L. Marie-Orleach, K. De Mulder, E. Berezikov, P. Ladurner, D. B. Vizoso, and L. Schärer. 2013. Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution 67: 3233-3242.

Johnston, M. O., B. Das, and W. R. Hoeh. 1998. Negative correlation between male allocation and rate of self-fertilization in a hermaphroditic animal. Proc. Natl. Acad. Sci. 95: 617-620.

Jousson, O., P. Bartoli, and J. Pawlowski. 2000. Cryptic speciation among intestinal parasites (Trematoda: Digenea) infecting sympatric host fishes (Sparidae). J. Evol. Biol. 13: 778–785.

Kasl, E.L., T. J. Fayton, W. F. Font, and C.D. Criscione. 2014. *Alloglossidium floridense* n. sp. (Digenea: Macroderoididae) from a Spring Run in North Central Florida. J. Parasitol. 100: 121-126.

Kozlowski, J. 1992. Optimal allocation of resources to growth and reproduction: Implications for age and size at maturity. Trends Ecol. Evol. 7: 15-19.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2015. lmerTest: Tests in Linear Mixed Effects Models. http://CRAN.R-project.org/package=lmerTest. Accessed 12 June 2015.

Lagrue, C. and R. Poulin. 2009. Life cycle abbreviation in trematode parasites and the developmental time hypothesis: is the clock ticking? J. Evo. Biol. 22: 1727-1738.

Lamont, M. E. 1921. Two new parasitic flatworms. Occasional Papers of the Museum of Zoology, University of Michigan 93: 1-4.

Lanfear, R., Calcott, B., Ho, S.Y.W., and S. Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Bio. Evo. 29: 1695-1701.

Lefebvre, F., and R. Poulin. 2005a. Progenesis in digenean trematodes: a taxonomic and synthetic overview of species reproducing in their second intermediate hosts. Parasitology 130: 587-605.

Lefebvre, F. and R. Poulin. 2005b. Life history constraints on the evolution of abbreviated life cycles in parasitic trematodes. J. Helminth. 79: 47-53.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lyons, J. 1989. Changes in the abundance of small littoral-zone fishes in Lake Mendota, Wisconsin. Canadian J. Zoo. 67: 2910-2916.

McAllister, C. T., H. W. Robison, and W. F. Font. 2011. Metacercaria of *Alloglossidium corti* (Digenea: Macroderoididae) from 3 species of crayfish (Decapoda: Cambaridae) in Arkansas and Oklahoma, U.S.A. Comp. Parasitol. 78: 382-386.

McAllister, C. T., W. F. Font, M. B. Connior, H. W. Robison, T. J. Fayton, N. G. Stokes, and C. D. Criscione. 2015. Trematode parasites (Digenea) of the Slender Madtom, *Noturus exilis*, and Black River Madtom, *Noturus maydeni*, (Siluriformes: Ictaluridae) from Arkansas, U.S.A. Comp. Parasitol. 82: 137-143.

McAllister, C. T., E. L. Kasl, H. W. Robison, M. B. Connior, W. F. Font, S. E. Trauth, and C. D. Criscione. New Host Records for *Alloglossidium progeneticum* (Digenea: Macroderoididae) in Crayfishes (Decapoda: Cambaridae) from Arkansas and Oklahoma, U.S.A. Comp. Parasitol. (*in press*).

McCoy, O. R. 1928. Life history studies on trematodes from Missouri. J. Parasitol. 14: 207-228.

Moran, N. 1994. Adaptation and constraint in the complex life cycles of animals. Annu. Rev. Ecol. Evol. Syst. 25: 573-600.

Morand, S., F. Robert, and V. A. Connors. 1995. Complexity in Parasite Life Cycles: Population Biology of Cestodes in Fish. J. Anim Ecol. 64: 256-264

Mueller, J. F. 1930. The trematode genus *Plagiorchis* in fishes. Transactions of the American Microscopical Society 49: 174–177.

Nollen, P. M. 1999. Mating behavior of *Echinostoma trivolvis* and *E. paraensei* in concurrent infections in hamsters. J. Helminthol. 73: 329-332.

Nuemann, M. P., and F. J. Vande Vusse. 1976. Two new species of *Alloglossidium* Simer 1929 (Trematoda: Macroderoididae) from Minnesota leeches. Journal of Parasitology 62: 556-559.

Olson, P. D., T. H. Cribb, V. V. Tkach, R. A. Bray, and D. T. J. Littlewood. 2003. Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int. J. Parasitol. 33: 733-755.

Parker, G. A., J. C. Chubb, M. A. Ball, and G. N. Roberts. 2003. Evolution of complex life cycles in helminth parasites. Nature 425: 480-484.

Parker, G. A., M. A. Ball, J. C. Chubb, K. Hammerschmidt, and M. Milinski. 2009. When should a trophically transmitted parasite manipulate its host? Evolution 63: 448-458.

Parker, G. A., M. A. Ball, and J. C. Chubb. 2015a. Evolution of complex life cycles in trophically transmitted helminths. I. Host incorporation and trophic ascent. J. Evol. Biol. 28: 267-291.

Parker, G. A., M. A. Ball, and J. C. Chubb. 2015b. Evolution of complex life cycles in trophically transmitted helminths. II. How do life-history stages adapt to their hosts. J. Evol. Biol. 28: 292-304.

Perkins, S. L., E. S. Martinsen, and B. G. Falk. 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138:1664-1674.

Pérez-Ponce de León, G. and S. A. Nadler. 2010. What we don't recognize can hurt us: A plea for awareness about cryptic species. J. Parasitol. 96: 453-464.

Poinar Jr., G. O., S. S. Schwarts, and G. Cameron. 1995. *Alloglossidium anomophagis* sp. n. (Trematoda: Plagiorchiidae) exhibiting progenesis in water fleas (Anomopoda: Daphniidae). Experientia 51: 388-390.

Poulin, R. and T. H. Cribb. 2002. Trematode life cycles: short is sweet? Trends Parasitol. 18:176-183.

Poulin, R. 2007. Evolutionary Ecology of Parasites. 2nd edition. Princeton University Press, Princeton, NJ.

Poulin, R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biol. Lett. 7:241-244.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rauch, G., M. Kalbe, T. B. H. Reusch. 2005. How a complex life cycle can improve a parasite's sex life. J. Evol. Biol. 18: 1069-1075.

Riggs, M. and M. J. Ulmer. 1983. Host-parasite relationships of helminth parasites in leeches of the genus *Haemopis*. I. Associations at the individual host level. Trans. Amer. Micro. Soc. 102(3): 213-226.

Roberts, L. S. and J. J. Janovy. 2009. Foundations of Parasitology. 8th edition. McGraw-Hill, New York.

Schärer, L. 2009. Tests of sex allocation theory in simultaneously hermaphroditic animals. Evolution 63: 1377-1405.

Schärer, L., M. Karlsson, M. Christen, and C. Wedekind. 2001. Size-dependent sex allocation in a simultaneous hermaphrodite parasite. J. Evol. Biol. 14: 55-67.

Schjørring, S. 2009. Sex allocation and mate choice of selfed and outcrossed *Schistocephalus solidus* (Cestoda). Behav. Ecol. 20: 644-650.

Schmidt, G. D., and K. Chaloupka. 1969. *Alloglossidium hirudicola* sp. n., a neotentic trematode (Plagiorchiidae) from leeches, *Haemopis* sp. J. Parasitol. 55: 1185-1186.

Simer, P. H. 1929. Fish trematodes from the lower Tallahatchie River. Amer. Mid. Nat. 11: 563–588.

Smythe, A. B. and W.F. Font. 2001. Phylogenetic analysis of *Alloglossidium* (Digenea: Macroderoididae) and related genera: Life-cycle evolution and taxonomic revision. J. Parasitol. 87: 386-391.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.

Strona, G. and S. Fattorini. 2014. Parasitic worms: How many really? Int. J. Parasitol. 44: 269-272.

Sullivan, J. J., and R. W. Heard III. 1969. *Macroderoides progeneticus* n. sp., a progenetic trematode (Digenea: Macroderoididae) from the antennary gland of the crayfish, *Procambarus spiculifer* (LeConte). T. Am. Microsc. Soc. 88: 304-308.

Tkach, V. V., T. J. Littlewood, P. D. Olson, J. M. Kinsella, and Z. Swiderski. 2003. Molecular phylogenetic analysis of the Microphalloidea Ward, 1901 (Trematoda: Digenea). System. Parasitol. 56: 1-15.

Tkach, V. V., and A. M. Mills. 2011. *Alloglossidium fonti* sp. nov. (Digenea: Macroderoididae) from black bullheads in Minnesota with molecular differentiation from congeners and resurrection of *Alloglossidium kenti*. Acta Parasitol. 56: 154-162.

Tkach V.V., Greiman S.E., and K. R. Steffes. 2013. *Alloglossidium demshini* sp. nov. (Digenea: Macroderoididae) from leeches in Minnesota. Acta Parasitol. 58: 434-440.

Trouve, S., J. J. Jourdane, F. Renaud, P. Durand, and S. Morand. 1999. Adaptive sex allocation in a simultaneous hermaphrodite. Evolution 53: 1599-1604.

Trouve, S., R. Renaud, P. Durand, and J. Jourdane. 1999. Reproductive and mate choice strategies in the hermaphroditic flatworm *Echinostoma caproni*. J. Hered. 90: 582-585.

Turner, H. M. and S. McKeever. 1993. *Alloglossoides dolandi* n. sp. (Trematoda: Macroderoididae) from the crayfish *Procambarus epicyrtus* in Georgia. J. Parasitol. 79: 353-355.

Weinzierl, R. P., K. Berthold, L. W. Beukeboom, and N. K. Michiels. 1998. Male allocation in the parthenogenetic hermaphrodite *Dugesia polychroa*. Evolution 52: 109-115.

Wilbur, H. M. 1980. Complex life cycles. Annu. Rev. Ecol. Evol. Syst. 11: 67-93.

Van Cleave, H. J., and J. F. Mueller. 1934. Parasites of Oneida Lake fishes. Part III. A biological and ecological survey of the worm parasites. Roosevelt Wildlife Annals 3: 163–334.

Vande Vusse, F. J. 1980. Revision of *Alloglossidium* Simer 1929 (Trematoda: Macroderoididae) and description of *A. microspinatum* sp. n. from a leech. J. Parasitol. 66: 667–670.

Vilas, R., C. D. Criscione, and M. S. Blouin. 2005. A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. Parasitology 131: 839-846.

APPENDIX

SUPPLEMENTAL MATERIAL

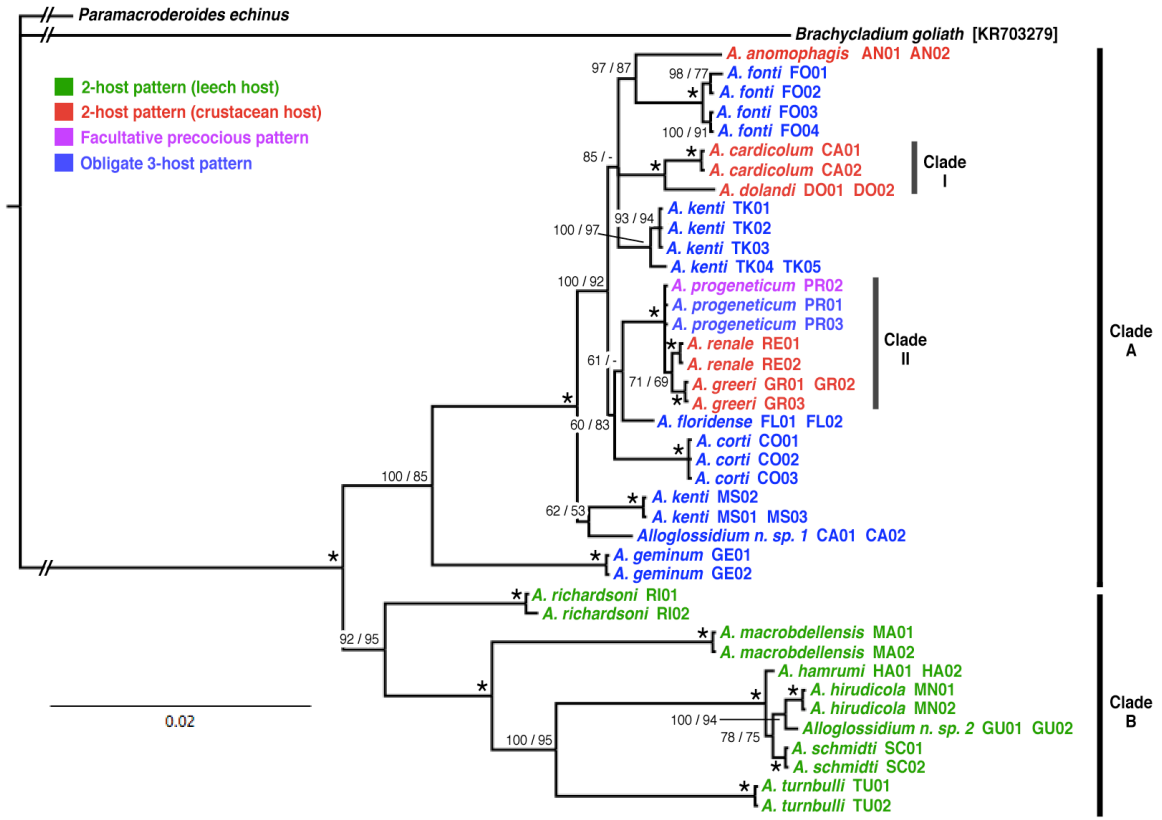

Figure S1. The 50% majority rule consensus tree obtained Bayesian analysis of the ribosomal-only (18s-ITS1-5.8s-ITS2-28s rDNA) dataset. Numbers at nodes correspond to the Bayesian posterior probability / RAxML maximum likelihood bootstrap support percentages; * denote nodal support values > 98% for both analyses. The outgroup taxa (*P. echinus* and *B. goliath*) have been collapsed and branch lengths shortened. Colors denote life cycle pattern (green = 2-host life cycle maturing in leech hosts; red = 2-host life cycle using crustacean final hosts; purple = facultative precocious; blue = obligate 3-host). Clades A, B, I, and II are designated based on congruence with the relationships reported from the concatenated dataset (see Fig. 4) and are explained further in the text.

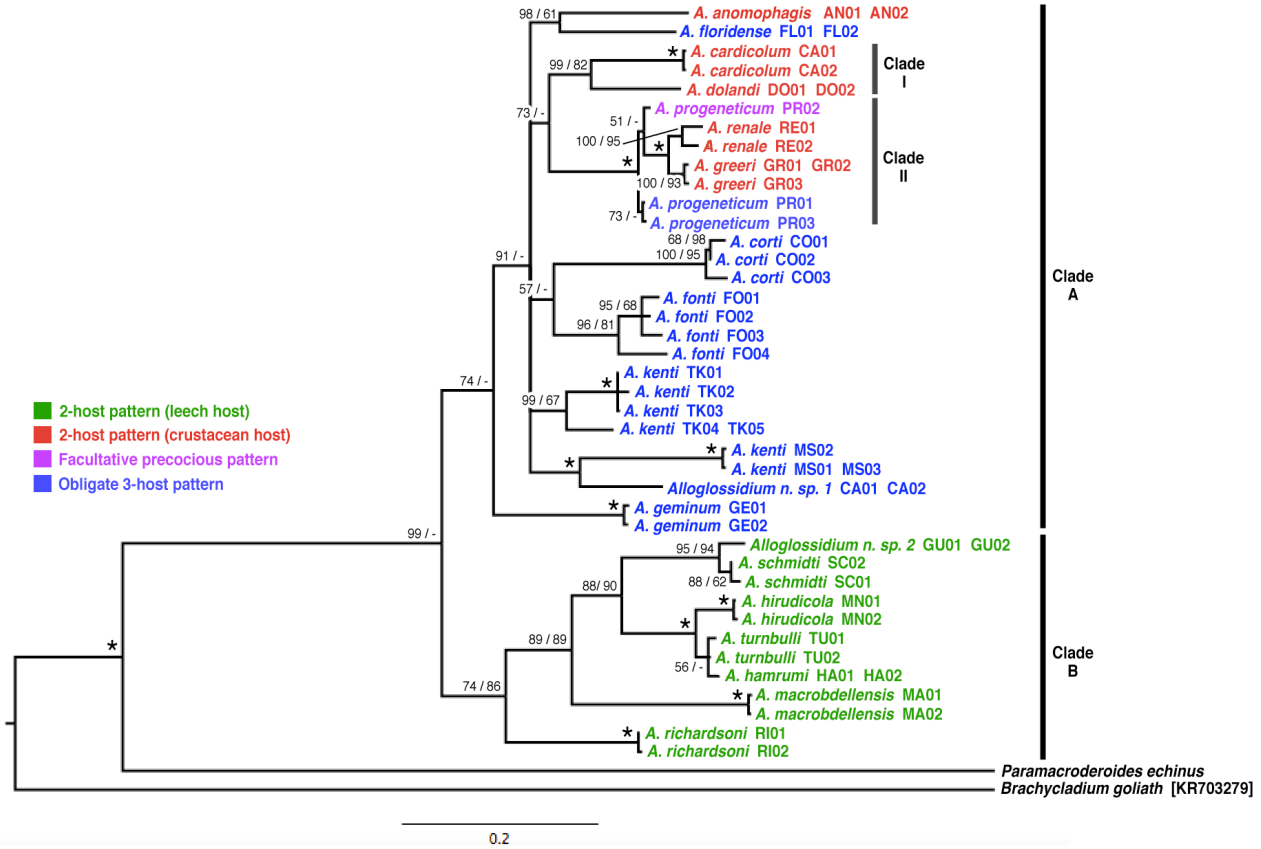

Figure S2. The 50% majority rule consensus tree obtained from the Bayesian analysis of the mitochondrial-only (ND1 mtDNA) dataset. The tree produced from the maximum likelihood (ML) RAxML analysis did not allow for resolution due to low support values, though the values > 50% are included on this reconstruction. Numbers at nodes correspond to the Bayesian posterior probability / maximum likelihood bootstrap support percentages; * denote nodal support values > 98% for both analyses. (-) denotes ML nodal support values of < 50%. Colors denote life cycle pattern (green = 2-host life cycle maturing in leech hosts; red = 2-host life cycle using crustacean final hosts; purple = facultative precocious; blue = obligate 3-host). Clades A, B, I, and II are designated based on congruence with the relationships reported from the concatenated dataset (see Fig. 4) and are explained further in the text.

Table S1. Pairwise % genetic distances (*p*-distance) of the ND1 mtDNA gene of all *Alloglossidium* specimens used for phylogenetic analyses (see Table 1). Individual (Ind) host IDs are provided.

No.	Ind.	Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	01/02	*A. anomophagis*																				
2	01	*A. cardicolum*	14.5																			
3	02	*A. cardicolum*	14.5	0.3																		
4	01/02	*A. dolandi*	13.6	11.4	11.4																	
5	02	*A. progeneticum*	12.9	12.1	12.1	11.8																
6	01	*A. progeneticum*	13.3	12.4	12.4	11.8	1.2															
7	03	*A. progeneticum*	13.3	12.4	12.4	11.8	1.3	0.4														
8	01	*A. renale*	13.9	13.3	13.3	14.1	4.4	5.0	5.2													
9	02	*A. renale*	13.9	13.6	13.6	12.9	4.1	4.7	4.7	3.0												
10	01/02	*A. greeri*	13.5	12.7	12.7	13.0	4.4	4.1	4.1	4.3	3.8											
11	03	*A. greeri*	13.2	12.7	12.7	13.2	4.3	4.0	4.1	4.1	4.0	0.7										
12	01/02	*A. floridense*	13.2	14.4	14.4	15.2	12.9	12.9	12.6	13.9	12.6	12.7	12.6									
13	01	*A. corti*	17.8	14.2	14.2	16.1	16.1	16.4	16.6	16.7	16.3	16.0	16.4	15.9								
114	02	*A. corti*	16.4	13.2	13.2	14.9	14.9	15.2	15.4	15.5	15.1	14.8	15.2	14.7	1.5							
15	03	*A. corti*	17.2	13.2	13.2	15.4	15.7	16.0	16.1	16.4	15.7	15.7	16.1	15.8	3.8	2.4						
16	01	*A. fonti*	12.9	14.1	14.1	13.5	9.8	9.9	10.1	11.4	11.4	10.9	10.8	12.9	14.1	12.9	13.5					
17	02	*A. fonti*	12.6	13.9	13.9	13.2	9.9	10.1	10.2	11.5	11.2	11.1	10.9	12.3	14.2	13.0	13.6	2.2				
18	03	*A. fonti*	13.0	13.5	13.5	13.8	10.4	10.9	10.7	12.1	12.4	12.0	12.0	13.3	15.4	14.2	14.2	5.9	5.5			
19	04	*A. fonti*	12.6	14.2	14.2	13.5	9.9	10.2	10.4	10.8	11.1	10.4	10.2	12.4	14.9	13.8	14.3	3.3	2.5	5.9		
20	01	*A. kenti* TK	13.0	11.4	11.4	12.0	11.1	11.4	11.4	11.5	11.4	12.4	12.6	13.0	14.2	13.0	13.3	11.4	11.4	11.4	11.7	
21	02	*A. kenti* TK	13.9	11.7	11.7	11.8	11.5	11.8	11.8	12.0	11.8	12.7	12.9	13.5	14.3	13.2	13.5	11.8	11.5	11.5	11.8	1.0
22	03	*A. kenti* TK	13.2	11.5	11.5	12.1	11.2	11.5	11.5	11.7	11.5	12.6	12.7	13.2	14.3	13.2	13.5	11.5	11.5	11.5	11.8	0.1
23	04/05	*A. kenti* TK	12.3	11.2	11.2	11.7	9.6	10.1	10.1	11.1	10.5	11.4	11.4	13.2	14.6	13.5	13.3	10.9	10.7	11.4	10.7	6.8
24	02	*A. kenti* MS	14.6	13.9	13.9	14.9	14.3	14.6	14.6	15.7	15.5	14.9	15.2	14.5	15.7	14.5	13.9	13.8	12.9	13.9	13.6	12.7
25	01/03	*A. kenti* MS	14.5	14.1	14.1	14.9	14.5	14.8	14.8	15.5	15.4	15.1	15.4	14.2	15.8	14.6	14.1	13.6	12.7	14.1	13.8	12.6
26	01/02	*Alloglossidium n. sp. 1*	14.3	13.8	13.8	13.9	12.7	12.7	13.2	13.3	13.2	13.8	13.6	14.5	16.3	15.1	15.5	12.4	11.5	13.0	12.1	10.8
27	01	*A. geminum*	15.6	15.3	15.3	15.6	13.9	13.6	13.6	15.5	14.2	14.7	15.0	15.0	16.7	15.8	15.2	12.7	11.8	13.2	12.9	12.6
28	02	*A. geminum*	15.1	14.6	14.6	15.5	13.8	13.5	13.6	15.2	14.6	14.9	14.9	14.7	16.3	15.4	15.1	12.7	11.7	13.2	12.7	12.9
29	01	*A. richardsoni*	17.0	16.0	16.0	16.9	15.8	16.0	16.0	17.3	17.0	16.1	16.4	16.7	18.5	17.3	17.2	14.9	13.9	14.9	14.9	14.8
30	02	*A. richardsoni*	16.7	15.7	15.7	16.6	15.5	15.7	15.7	17.0	16.7	15.8	16.1	16.4	18.2	17.0	16.9	14.6	13.6	14.6	14.6	14.5
31	01	*A. macrobdellensis*	18.8	16.7	16.7	15.5	16.4	16.3	16.4	17.2	17.5	16.6	16.9	17.6	20.1	18.9	18.9	15.1	14.6	15.7	15.4	15.8
32	02	*A. macrobdellensis*	18.6	16.6	16.6	15.7	16.3	16.1	16.3	16.9	17.3	16.4	16.7	17.5	20.0	18.8	18.8	14.8	14.3	15.4	15.1	15.8
33	01/02	*A. hamrumi*	18.0	18.0	18.0	17.9	17.9	17.5	17.2	17.2	17.2	17.3	17.8	17.9	21.3	20.0	20.1	16.1	15.8	16.6	16.0	16.0
34	01	*A. hirudicola*	18.6	17.8	17.8	18.6	17.9	18.2	18.0	17.8	17.6	17.6	17.8	17.8	20.6	19.2	19.5	16.9	16.3	17.0	16.3	16.4
35	02	*A. hirudicola*	18.8	17.8	17.8	18.5	17.9	18.2	17.9	17.8	17.3	17.3	17.8	17.6	20.4	19.1	19.4	17.0	16.4	17.0	16.4	16.3
36	01	*A. schmidti*	19.0	18.0	18.2	18.7	17.9	18.2	18.4	18.5	18.4	17.5	17.7	18.2	20.7	19.7	19.9	17.1	16.2	16.9	16.9	16.9
37	02	*A. schmidti*	19.2	17.9	18.0	18.5	17.7	18.0	18.0	18.7	18.2	17.1	17.5	18.2	20.9	19.5	19.7	17.2	16.4	16.6	16.7	16.7
38	01/02	*Alloglossidium n. sp. 2*	18.0	16.9	16.9	17.6	16.9	16.9	16.9	17.3	17.2	15.8	16.4	17.8	19.8	18.6	18.6	15.7	14.9	14.6	15.2	15.4
39	01	*A. turnbulli*	18.2	17.8	17.8	17.9	17.5	17.0	16.7	17.0	16.9	16.7	17.2	17.3	20.7	19.4	19.5	16.3	16.0	17.2	16.4	15.5
40	02	*A. turnbulli*	18.6	18.3	18.3	18.6	17.8	17.3	17.0	17.3	17.5	17.0	17.5	17.6	21.4	20.1	20.3	16.6	16.0	17.2	16.4	16.0

Table S1. Continued

No.	Ind.	Taxa	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	01/02	*A. anomophagis*																				
2	01	*A. cardicolum*																				
3	02	*A. cardicolum*																				
4	01/02	*A. dolandi*																				
5	02	*A. progeneticum*																				
6	01	*A. progeneticum*																				
7	03	*A. progeneticum*																				
8	01	*A. renale*																				
9	02	*A. renale*																				
10	01/02	*A. greeri*																				
11	03	*A. greeri*																				
12	01/02	*A. floridense*																				
13	01	*A. corti*																				
114	02	*A. corti*																				
15	03	*A. corti*																				
16	01	*A. fonti*																				
17	02	*A. fonti*																				
18	03	*A. fonti*																				
19	04	*A. fonti*																				
20	01	*A. kenti* TK																				
21	02	*A. kenti* TK																				
22	03	*A. kenti* TK	1.2																			
23	04/05	*A. kenti* TK	7.7	7.0																		
24	02	*A. kenti* MS	12.9	12.9	12.9																	
25	01/03	*A. kenti* MS	12.7	12.7	12.7	0.4																
26	01/02	*Alloglossidium n. sp. 1*	11.1	10.9	10.8	12.0	11.8															
27	01	*A. geminum*	12.7	12.7	12.4	13.3	13.2	15.0														
28	02	*A. geminum*	13.2	13.0	12.4	13.5	13.5	14.6	0.8													
29	01	*A. richardsoni*	14.6	14.9	14.2	15.1	15.5	15.5	15.6	15.4												
30	02	*A. richardsoni*	14.3	14.6	13.9	14.8	15.2	15.2	15.3	15.1	0.3											
31	01	*A. macrobdellensis*	16.3	15.7	17.0	16.6	16.9	16.9	17.6	17.3	16.6	16.3										
32	02	*A. macrobdellensis*	16.3	15.7	16.9	16.3	16.6	16.7	17.6	17.3	16.4	16.1	0.4									
33	01/02	*A. hamrumi*	16.3	16.1	16.9	16.4	16.6	16.1	17.3	17.6	15.4	15.1	15.5	15.5								
34	01	*A. hirudicola*	16.4	16.6	17.9	17.5	17.3	16.4	17.8	18.0	15.5	15.2	15.1	15.4	4.9							
35	02	*A. hirudicola*	16.3	16.4	17.6	17.3	17.2	16.6	17.8	18.3	15.4	15.1	15.2	15.5	4.7	0.4						
36	01	*A. schmidti*	16.9	17.1	17.7	16.1	15.7	15.1	18.5	18.4	15.7	15.4	15.9	16.2	12.1	11.3	11.3					
37	02	*A. schmidti*	16.7	16.9	17.5	15.6	15.6	15.2	18.4	18.5	15.2	14.9	15.7	16.1	11.9	11.4	11.1	0.8				
38	01/02	*Alloglossidium n. sp. 2*	15.5	15.5	16.0	15.8	16.0	14.3	16.8	16.6	14.8	14.5	14.3	14.6	12.6	12.7	12.3	3.8	3.0			
39	01	*A. turnbulli*	15.8	15.7	16.7	16.4	16.6	15.7	16.7	17.2	15.2	14.9	15.4	15.7	1.6	4.7	4.6	11.4	11.3	12.1		
40	02	*A. turnbulli*	16.3	16.1	17.2	16.9	17.0	16.4	17.3	17.6	14.9	14.6	15.2	15.5	2.4	4.9	4.7	11.9	11.8	12.6	1.0	

Table S2. Locality, host type, life stage (M=metacercariae, E=encysted adult (i.e. a gravid worm), A=adult), prevalence (%), and mean intensity (±1 standard deviation, range in parentheses) for *Alloglossidium progeneticum* collected across sites spanning the Southeastern United States (see Figure 10). Also given are the number of worms sequenced and haplotype IDs (see Figure 12).

Locality	Lat (N)	Long (W)	Collection Date	Host	Stage	Prevalence	Intensity	No. Sequenced	Haplotypes
Alabama:									
Black Oak	34.41722	86.04027	05/2013	*Ameiurus natalis*^	A	8/10 (80)	6 ± 6.8 (1-21)	2	Hp01
Franklin Creek	30.47093	88.3859	12/2013	*Procambarus clarkii*	M	3/6 (50)	1.7 ± 0.6 (1-2)	1	Hp01
Minky Creek	34.439	86.1761	05/2013	*Procambarus sp.*	M	5/21 (24)	2.8 ± 2.2 (1-6)	1	Hp02
Arkansas:									
Calhoun Co.	33.513888	92.371588	02/2013	*Procambarus*	M	1/1 (100)	1	1	Hp02
Crooked Creek	36.245207	92.715611	07/2013	*Ameiurus natalis*^	A	7/7 (100)	27.7 ± 28.1 (1-76)	8	Hp02, Hp03
			07/2012-	*Noturus exilis*^	A	3/31 (10)	2.3 ±1.2 (1-3)	1	Hp02
			07/2011	*Orconectes neglectus*	M	1/1 (100)	1	1	Hp02
Creek & Hwy	36.44023	91.48923	07/2013	*Ameiurus natalis*^	A	1/2 (50)	20	2	Hp02, Hp04
Nix Creek	33.444116	94.016049	07-	*Ameiurus natalis*^	A	15/23 (65)	2.6 ± 3.8 (1-15)	3	Hp01, Hp05
			03-	*Procambarus acutus*	M	16/20 (80)	13.1 ± 9.3 (2-33)	9	Hp01, Hp05,
			03-	*Procambarus*	M	13/15 (87)	18.2 ± 12.9 (6-43)	6	Hp01, Hp09,
Water Creek	36.0525	92.595111	08/2012	*Noturus exilis*^	A	1/1 (100)	2	2	Hp03, Hp11
			05/2012	*Orconectes p. palmeri*	M	1/1 (100)	10	2	Hp11, Hp12
Georgia:									
Big Indian	33.37632	83.47143	05/2013-	*Ameiurus brunneus*^	A	6/15 (40)	3.3 ± 2.3 (1-7)	5	Hp13
			05/2012-	*Procambarus spiculifer*	E	56/62 (90)	11.4 ± 9.7 (1-41)	5	Hp13, Hp14
Calls Creek*	33.8885	83.38223	05/2013-	*Ameiurus brunneus*^	A	7/63 (11)	4.7 ± 4.4 (1-12)	5	Hp13, Hp14
			05/2012-	*Procambarus spiculifer*	E	58/62 (94)	27.5 ± 22.0 (1-104)	5	Hp14
Jackson Branch	32.6145	81.73095	07/2011	*Procambarus acutus*	M	7/10 (70)	3.3 ± 2.9 (1-8)	2	Hp15, Hp16
Mill Creek	34.71567	85.42562	05/2013	*Ameiurus natalis*^	A	2/2 (100)	2.5 ± 2.1 (1-4)	1	Hp17
			05/2013	*Orconectes erichsonianus*	M	1/6 (17)	2	1	Hp17
Richland Creek	32.69762	84.07288	05/2014	*Ameiurus brunneus*^	A	1/5 (20)	3	2	Hp18
			05/2013-	*Ameiurus natalis*^	A	25/30 (83)	6.5 ± 7.6 (1-19)	2	Hp18
			05/2014	*Noturus lepthocanthus*^	A	5/6 (83)	5.2 ± 3.8 (1-10)	2	Hp18
			05/2013-	*Procambarus spiculifer*	E	28/34 (82)	4.1 ± 3.6 (1-16)	4	Hp18

* Denotes type locality ^ Denotes fish host

Table S2. Continued.

Locality	Lat (N)	Long (W)	Collection Date	Host	Stage	Prevalence	Intensity	No. Sequenced	Haplotypes
Louisiana:									
Chappepeela Cr.	30.6823	90.31658	10/13-03/14	*Ameiurus natalis*^	A	3/3 (100)	3.0 ± 1.0 (2-4)	4	Hp17, Hp19,
			03/2012	*Cambarellus shufeldtii*	M	4/6 (67)	1.5 ± 0.6 (1-2)	2	Hp21, Hp22
			03/2012	*Procambarus clarkii*	M	1/22 (5)	5	1	Hp17
			03/2012	*Procambarus vioscai*	M	2/5 (40)	1.5 ± 0.7 (1-2)	3	Hp17, Hp19
Cr. & LA-169	32.467083	93.98175	03/2013	*Ameiurus natalis*^	M	2/2 (100)	17.5 ± 17.7 (5-30)	1	Hp23
Sanchez Rd	29.85312	90.67825	03/2012	*Cambarellus shufeldtii*	M	1/5 (20)	1	1	Hp01
Head of Island	30.26285	91.556617	03/2010	*Procambarus clarkii*	M	1/2 (50)	1	1	Hp01
Oklahoma:									
Boktuklo Creek	34.074691	94.852014	07/2012	*Procambarus acutus*	M	1/1 (100)	1	1	Hp24
Eagle Fork Creek	34.450583	94.663074	07/2011	*Orconectes p. longimanus*	M	3/7 (43)	1.3 ± 0.6 (1-2)	1	Hp25
Mt. Fork River	34.138315	94.687941	08/2011	*Orconectes p. longimanus*	M	8/18 (44)	2.1 ± 1.1 (1-5)	1	Hp26
Gaines Creek	34.830351	95.311026	08/2011	*Orconectes p. longimanus*	M	1/4 (25)	9	2	Hp27, Hp28
Glover River	34.097954	94.902678	07/2011	*Orconectes p. longimanus*	M	3/8 (38)	1.7 ± 1.2 (1-3)	1	Hp29
Lukfata Creek	33.968825	94.766173	10/2011	*Ameiurus natalis*^	A	1/10 (10)	2	1	Hp01
			06-10/2011	*Orconectes p. longimanus*	M	14/21 (67)	2.2 ± 1.4 (1-5)	2	Hp30, Hp31
Salt Creek	33.920267	94.803144	07/2011	*Orconectes p. longimanus*	M	5/10 (50)	1.6 ± 0.5 (1-2)	1	Hp32
Yanubbee Creek	34.064371	94.73986	07/2011	*Orconectes p. longimanus*	M	5/6 (83)	10.2 ± 8.6 (4-25)	2	Hp27, Hp33
Yashau Creek	34.019156	94.75496	06/2011	*Orconectes p. longimanus*	M	13/18 (72)	4.0 ± 3.1 (1-10)	5	Hp27, Hp34
Texas:									
Gus Engeling	31.74468	95.8896	05/2013	*Ameiurus natalis*^	A	38/39 (97)	18.5 ± 15.2 (4-78)	3	Hp01, Hp35
Wildlife			05/2013	*Noturus gyrinus*^	A	1/1 (100)	8	1	Hp01
Management			03/2012-13	*Procambarus acutus*	M	19/36 (53)	25.9 ± 27.0 (1-102)	3	Hp01, Hp35
Area			04-05/2013	*Procambarus clarkii*	M	11/18 (61)	11.1 ± 19.4 (2-54)	2	Hp01, Hp05
			05/2013	*Procambarus kensleyi*	M	1/3 (33)	14	1	Hp36
Whites Head Cr.	30.56752	96.37007	03/2013	*Ameiurus natalis*^	A	3/5 (60)	1.7 ± 0.6 (1-2)	2	Hp29, Hp37

* Denotes type locality ^ Denotes fish host

Table S3. Comparative morphometrics with mean values ± standard deviation (range) of adult *Alloglossidium progeneticum* specimens across the Southeastern United States. Populations were pooled into facultative precocious (Calls Creek, GA (N = 10); Big Indian Creek, GA (N = 11); Richland Creek, GA (N = 12)) or obligate 3-host (Gus Engeling Wildlife Management Area, TX (N = 10); Chappapeela Creek, LA (N = 10); Crooked Creek, AR (N = 9)) populations based on life history patterns. Note there is no correction for body size in these measurements (see text for appropriate statistical analyses). Morphologically, the extent of the vitellaria can be used as a diagnostic characteristic when compared with the 5 other *Alloglossidium* species using fish definitive hosts (vitellaria extend from just below cecal bifurcation to the anterior margin of the posterior testis; Fig. 11; see Kasl et al. 2014).

Measurements (μm)	Facultative Precocious Populations (n = 33)	Obligate 3-host Populations (n = 29)
Body length	1446.1 ± 439.4 (889-2300)	1724.1 ± 496.9 (825-2500)
Body width at ventral sucker	320.5 ± 131.3 (202-950)	314.5 ± 81.3 (175-465)
Forebody	420.8 ± 95.4 (242-606)	458.7 ± 115.4 (278-717)
Oral sucker length	120.9 ± 28.8 (61-182)	120.8 ± 23.8 (81-182)
Oral sucker width	139.8 ± 28.2 (91-202)	139.5 ± 33.5 (81-202)
Prepharynx length	82.6 ± 38.9 (20-152)	79.4 ± 28.6 (35-182)
Pharynx length	62.0 ± 13.1 (40-91)	65.7 ± 19.4 (35-111)
Pharynx width	60.4 ± 17.0 (30-91)	64.1 ± 21.0 (30-111)
Esophagus	32.9 ± 13.6 (15-71)	35.2 ± 16.1 (15-101)
Caecal bifurcation to anterior end	311.4 ± 73.6 (172-465)	324.9 ± 89.4 (172-534.5)
Ventral sucker length	112.6 ± 29.2 (71-162)	117.2 ± 22.6 (76-157)
Ventral sucker width	114.8 ± 31.1 (61-182)	119.5 ± 28.6 (76-172)
Cirrus sac length	161.6 ± 61.7 (56-303)	219.6 ± 66.9 (126-384)
Cirrus sac width	38.0 ± 16.8 (10-71)	49.6 ± 19.5 (25-91)
Genital pore to anterior end	408.5 ± 104.4 (238-667)	440.1 ± 111.9 (258-687)
Vitelline field to anterior end	347.1 ± 87.4 (141-505)	375.2 ± 92.4 (222-606)
Vitelline field to posterior end	432.7 ± 204.1 (152-909)	560.8 ± 189.2 (207-848)
Ovary length	139.8 ± 32.0 (91-212)	142.7 ± 35.4 (86-202)
Ovary width	128.5 ± 36.1 (81-253)	130.7 ± 26.7 (76-177)
Ovary to ventral sucker	28.4 ± 28.1 (0-111)	40.2 ± 32.9 (0-141)
Anterior testis length	116.8 ± 48.9 (56-242)	178.7 ± 57.1 (86-268)
Anterior testis width	115.4 ± 42.4 (61-202)	183.9 ± 55.3 (96-328)
Posterior testis length	130.6 ± 58.9 (51-283)	202.3 ± 50.6 (126-318)
Posterior testis width	127.5 ± 45.2 (71-253)	194.3 ± 45.4 (96-263)
End of caeca to posterior end	153.8 ± 75.2 (35-283)	151.0 ± 58.4 (51-303)
Egg length	21.9 ± 1.7 (18-25)	21.5 ± 2.2 (14-24)
Egg width	11.2 ± 13.0 (9-13)	10.5 ± 1.3 (8-12)

Table S4. *Alloglossidium progeneticum*: variable factor loadings, factor eigenvalues, and percent total variance accounted for by each factor from the Varimax rotated correlation matrix of all sampled worms.

	Varimax rotated loading matrix	
	Factor 1	Factor 2
Ovary length	0.353	**0.829**
Ovary width	0.224	**0.903**
Anterior Testes length	**0.919**	0.287
Anterior Testes width	**0.904**	0.303
Posterior Testes length	**0.931**	0.27
Posterior Testes width	**0.898**	0.32
Eigenvalues	3.511	1.851
Percent of total variance explained by the factor	58.524	30.853

* Bold denotes loadings where variable loaded onto factor

Table S5. Infection intensity linear mixed model analyses using restricted maximum likelihood (REML) in R. The interaction between the fixed effect of intensity of infection (high vs. low) and covariate of worm body length (μm) was not significant for any of the dependent variables (Sex allocation: $F_{1,41} = 0.75$, $P = 0.39$; Female summated score: $F_{1,41} = 1.85$, $P = 0.18$; Male summated score: $F_{1,41} = 2.94$, $P = 0.09$). Thus, these interactions were pooled for the analyses reported below. *P* values for were calculated using the Satterthwaite approximation (fixed effects) or likelihood ratio test (random effects) using the lmerTest package. Significant effects are bolded.

A) Sex Allocation

lmer(Sex Allocation ~ Intensity+Body Size + (1 | Population) + (1 | Population:Intensity) + (1 | Population:Intensity:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Intensity	9.59	1	42	**0.003**
Body Size	13.58	1	42	**< 0.001**
Random Effects		**Variance**		
Population		0		1
Population:Intensity		0		1
Population:Intensity:Host		0		1
Residual		4.02×10^{-4}		
Total		4.02×10^{-4}		

B) Female Summated Score

lmer(Female SumScore ~ Intensity+Body Size + (1 | Population) + (1 | Population:Intensity) + (1 | Population:Intensity:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Intensity	2.57	1	42	0.12
Body Size	63.04	1	42	**< 0.001**
Random Effects		**Variance**		
Population		0		1
Population:Intensity		0		1
Population:Intensity:Host		0		1
Residual		1344		
Total		1344		

C) Male Summated Score

lmer(Male SumScore ~ Intensity+Body Size + (1 | Population) + (1 | Population:Intensity) + (1 | Population:Intensity:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Intensity	0.38	1	42	0.54
Body Size	176.72	1	42	**< 0.001**
Random Effects		**Variance**		
Population		0		1
Population:Intensity		0		1
Population:Intensity:Host		0		1
Residual		7208		
Total		7208		

Table S6. Mating opportunity linear mixed model analyses using restricted maximum likelihood (REML) in R. The interaction between the fixed effect of mating opportunity (forced self-fertilization in crayfish vs. opportunity to outcross in fish) and covariate of worm body length (μm) was not significant for the female summated score ($F_{1,54} = 1.54$, $P = 0.22$); and thus was pooled in the analysis reported below. "Host" refers to the host individual while "host type" refers to the host species (crayfish vs. fish). *P* values for were calculated using the Satterthwaite approximation (fixed effects) or likelihood ratio test (random effects) using the lmerTest package. Significant effects are bolded.

A) Sex Allocation

*lmer(Sex Allocation ~ Host Type*Body Size + (1 | Population) + (1 | Population:Host Type) + (1 | Population:Host Type:Host))*

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Host Type	1.16	1	19	0.29
Body Size	3.72	1	55	0.06
Host Type*Body Size	11.25	1	55	**0.001**
Random Effects		**Variance**		
Population		1.81×10^{-4}		0.7
Population:Host Type		3.90×10^{-4}		0.1
Population:Host Type:Host		1.75×10^{-4}		0.3
Residual		8.76×10^{-4}		
Total		1.62×10^{-3}		

B) Female Summated Score

lmer(Female SumScore ~ Host Type+Body Size + (1 | Population) + (1 | Population:Host Type) + (1 | Population:Host Type:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Host Type	0.99	1	19	0.33
Body Size	71.27	1	56	**< 0.001**
Random Effects		**Variance**		
Population		0		1
Population:Host Type		0		1
Population:Host Type:Host		148.4		0.4
Residual		1287.3		
Total		1435.7		

C) Male Summated Score

*lmer(Male SumScore ~ Host Type*Body Size + (1 | Population) + (1 | Population:Host Type) + (1 | Population:Host Type:Host))*

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Host Type	6.99	1	9	**0.03**
Body Size	132.05	1	56	**< 0.001**
Host Type*Body Size	39.19	1	56	**< 0.001**
Random Effects		**Variance**		
Population		150.5		0.96
Population:Host Type		3627		**0.001**
Population:Host Type:Host		8.42×10^{-12}		1
Residual		3601		
Total		7378.5		

Table S7. Life history linear mixed model analyses using restricted maximum likelihood (REML) in R. The interaction between the fixed effect of life history (facultative precocious vs. obligate 3-host) and covariate of worm body size (μm) was not significant for either female ($F_{1,23} = 0.46$, $P = 0.51$); or male ($F_{1,55} = 0.59$, $P = 0.44$); summated scores and thus was pooled the analyses reported below. P values for were calculated using the Satterthwaite approximation (fixed effects) or likelihood ratio test (random effects) using the lmerTest package. Significant effects are bolded.

A) Sex Allocation

*lmer(Sex Allocation ~ Life History*Body Size + (1 | Life History:Population) + (1 | Life History:Population:Host))*

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Life History	17.5	1	19	**<0.001**
Body Size	11.43	1	56	**0.001**
Life History*Body Size	5.17	1	56	**0.03**
Random Effects		**Variance**		
Life History:Population		6.17×10^{-4}		**0.007**
Life History:Population:Host		2.08×10^{-4}		**0.04**
Residual		4.59×10^{-4}		
Total		1.28×10^{-3}		

B) Female Summated Score

lmer(Female SumScore ~ Life History+Body Size + (1 | Life History:Population) + (1 | Life History:Population:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Life History	2.74	1	4	0.18
Body Size	70.4	1	26	**< 0.001**
Random Effects		**Variance**		
Life History:Population		42.49		0.8
Life History:Population:Host		350.42		0.07
Residual		1044.19		
Total		1437.1		

C) Male Summated Score

lmer(Male SumScore ~ Life History+Body Size + (1 | Life History:Population) + (1 | Life History:Population:Host))

Fixed Effect	*F*-value	df num.	df den.	*P-value*
Life History	10.87	1	4	**0.03**
Body Size	150.4	1	58	**< 0.001**
Random Effects		**Variance**		
Life History:Population		3875		**0.008**
Life History:Population:Host		1091		0.23
Residual		3885		
Total		8851		